PART I ANAESTHESIA REFRESHER COURSE
2017

DEPARTMENT OF ANAESTHESIA
& PERIOPERATIVE MEDICINE
UNIVERSITY OF CAPE TOWN

ACKNOWLEDGEMENTS

The committee wishes to thank the following companies for their generous assistance:

Teleflex Medical (Pty) Ltd

Registration and financial assistance

B. Braun Medical (Pty) Limited
Edwards Lifesciences (Pty) Ltd
Fresenius Kabi South Africa (Pty) Ltd
Intersurgical (RSA) (Pty) Ltd
Karl Storz Endoscopy South Africa (Pty) Ltd
Medhold Medical (Pty) Ltd
Medication Management Solutions
Safeline Pharmaceuticals (Pty) Ltd
SSEM Mthembu Medical (Pty) Ltd

financial assistance

COMMITTEE

Prof. J Swanevelder	-	*Head of Department*
Dr M Nejthardt	-	*Chairman / Programme / Trade*
Prof. R Dyer	-	*Programme*
Dr R Llewellyn	-	*Secretary / Treasurer*
Dr O Porrill	-	*Venue / Catering*
Dr R Nieuwveld	-	*Website / Printing*
Dr D van Dyk	-	*Programme / Book compilation*
Dr R Haylett	-	*Book compilation*
Mrs Z Carlse	-	*Secretary*

This course is held under the auspices of, and supported by, the

Colleges of Medicine of South Africa

Index of lectures

The "Big Five" in Physiology and Physics:
Laws-Equations-Principles

Professor Justiaan Swanevelder

Head of Department
UCT Dept of Anaesthesia & Perioperative Medicine

Several physics equations are used to explain concepts of medical physics and human physiology, and are regularly applied in anaesthesia. Many of these concepts are being used on a daily basis in clinical practice without even thinking about the theory behind them. Philosophy gave birth to education and science! Most of the greatest scientists were also famous philosophers and started off with philosophical ideas. Several famous Philosopher-Physicists-Physiologists have their names connected to one or more "Law-Equation-Principle" in the medical field.

Laws-Equations-Principles used every day & that should be understood in preparation for the Primary and Final FCA Exam, and for the rest of the anaesthetist/clinician's working life, include:

1. *Starling and compliance* (cardiac function/mechanics, lung physiology, renal function)
2. *La Place, Ohm* (wall tension, preload, resistance, afterload, oxygen consumption)
3. *Bernoulli* (flow hydraulics, ultrasound-Doppler, pressure drop/flow across narrowing)
4. *Fick, Shunt and Dead Space* (cardiac, respiratory - diffusion, filters, oxygen supply and demand, monitoring)
5. *Hagen-Poiseuille* (flow through airways, tubes and tubing/vessels)

Also important, but not covered in this lecture:

6. *Boyle* (Body plethysmography, "syringe, antibiotics, air and vial")
7. *Bohr* (Haldane and Bohr Effects - O_2-Hb Dissociation Curve, Dead Space Ventilation)
8. *Henderson-Hasselbach* (Blood gas interpretation)
9. *Coanda* (air/fluid flow pattern)
10. *Venturi* (gas entrailment, flow pattern)
11. *Stewart-Hamilton* (Cardiac Output by thermodilution)
12. *Alveolar Gas Equation* (Altitude)

etc……

1. Starling (Law - Curve - Mechanism - Equation - Resistor – Forces - Principle)

Frank–Starling curve

The function of the heart as a pump is based on cardiac muscle. The contractile properties of heart muscle not only provide the engine to drive the cardiac pump but also give the heart an intrinsic ability to adapt its performance to a continually varying venous return. The mechanism underlying this adaptive ability is the Frank–Starling relationship.

Frank curve

Frank demonstrated in isolated muscle fiber preparations that the tension developed on contraction was dependent on the initial length of the fiber. As initial length increased from resting value, the tension developed during contraction increased, and reached a maximum. Above this, the tension declined as the sarcomeres became overextended.

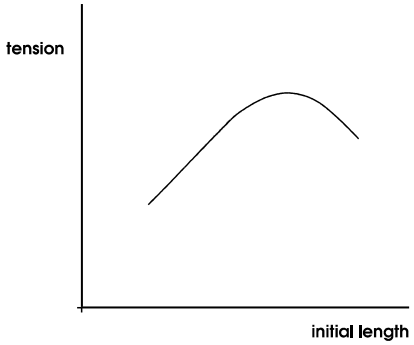


Figure: Frank curve for isolated muscle

Starling curve

The above property of isolated cardiac muscle fibers can be applied to the muscle fibers in the walls of an intact ventricle, where the length of muscle fiber is related to the volume in the ventricle. In this case tension per unit cross section (ventricular wall stress T), developed in the wall during contraction, is dependent on end-diastolic volume. Laplace's law relates the wall stress to internal pressure in an elastic sphere; thus, the Frank relationship for an isolated muscle fiber translates into a relationship between intraventricular pressure and EDV during isovolumetric contraction. Effectively, the greater the ventricular filling volume, the stronger the contraction of the ventricle – a mechanism that gives the intact heart its built-in ability to adjust to varying levels of venous return. Starling confirmed in ejecting mammalian hearts that with a constant aortic pressure, an increase in EDV produces a more forceful contraction and an increase in SV.

Frank–Starling Law/Frank-Starling Mechanism

In the intact heart, a ventricular function curve (Frank–Starling curve) can be plotted to demonstrate the ability of the ventricle to vary its mechanical output according to its filling volumes. An index of mechanical output (such as SV) can be plotted against a measure of filling pressure (such as CVP).

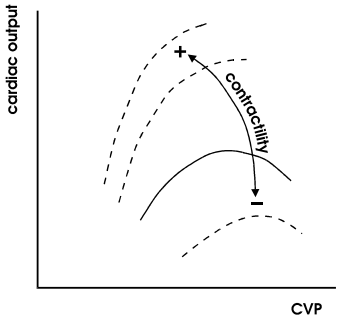


Figure: Ventricular function curve

Frank–Starling curve and cardiac failure

A normal ventricle never fills to an EDV that would place it on the descending limb of the Frank–Starling curve. This is because of decreased compliance of the ventricle that occurs at high filling pressures. Sarcomere length at optimum filling pressures (about 12 mmHg) is 2.2 μm; however, even if filling pressures are increased four-fold (>50 mmHg) sarcomere length will not increase beyond about 2.6 μm.

If the heart becomes pathologically dilated as in cardiac failure, ventricular function may shift to the descending portion of the Frank–Starling curve, and cardiac decompensation ensues. Cardiac function can also deteriorate when factors such as hypoxia, acidosis or β blockers shift the Frank–Starling curve down and to the right, depressing cardiac performance. Alternatively, other factors such as endogenous catecholamine or inotropes can shift the Frank–Starling curve upwards and to the left, enhancing cardiac performance.

Force–velocity curve for cardiac muscle

Starling not only investigated the sarcomere tension–length relationship, but also looked at the interaction between muscle force and velocity. The force–velocity curve demonstrates that the force generated and the velocity of muscle shortening is inversely related. Changes in preload and contractility will influence this relationship by shifting the force–velocity curve.

Starling was a wise man and thought about many other physiological principles, e.g.

He showed that there are opposing forces across any capillary wall, e.g. lung or kidney. Water is forced out through the pores in the wall by hydrostatic pressure, and driven in by the osmotic pressure of plasma proteins within the capillary. These opposing forces approximately balance and the concept is known as **Starling's Principle**.

Net fluid out = $K [(P_c – P_i) – \sigma (\pi_c – \pi_i)]$

Transmural pressure, P = hydrostatic pressure, π **=** osmotic pressure, K = filtration coefficient, σ = reflection coefficient

The hot topic of today is the "glycocalyx" of the vascular (alveolar/renal) membrane, which can conceptually be incorporated into K and **σ.** This still leaves the Starling Principle in situ!

The **Starling Resistor** was used in his isolated heart preparations that led to the description of "Frank-Starling Law of the Heart". It consisted of an elastic fluid-filled collapsible-tube mounted inside a chamber filled with air.

The static pressure inside the chamber was used to control the degree of collapse of the tube, so providing a variable resistor. This resistance was used to simulate TPR, or total peripheral (vascular) resistance. It is also used to describe pulmonary blood flow through the different "West Zones" and the influence of gravity and pressure.

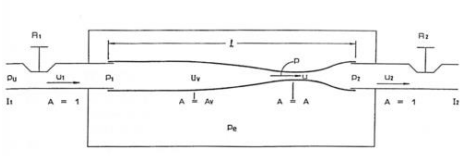


Figure: Starling Resistor

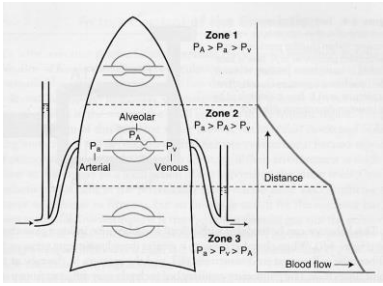


Figure: West Zones of the lung

From: Respiratory Physiology-the essentials, JB West, Ninth Edition

2. La Place Law

T=Pr/h

Where T = Wall Tension, P = Intracavity Pressure, r = Radius of Ventricular Cavity, h = Ventricular Wall Thickness

La Place's Law is often used in the setting of cardiovascular physiology as basic definitions like preload and afterload. It is also used to explain cardiac vulnerability in the setting of pathologies like ischemic heart disease, dilated cardiomyopathy, or valvular heart disease. The biggest physiological influence on oxygen consumption of the myocardium is the tension in its wall, which according to the equation is influenced directly by intraventricular pressure, cavity radius, and inversely by the wall thickness.

Preload

In clinical circumstances, 'preload' remains loosely defined and has become synonymous with a range of parameters including CVP, venous return and pulmonary capillary wedge pressure. A strict definition for preload can be obtained from the Frank relationship between muscle fiber length and developed tension. Here preload is the initial length of the muscle fiber before contraction. In the intact ventricle the preload would, therefore, be equivalent to the end-diastolic volume, since the pre-systolic length of the myocardial fibers will be directly related to EDV.

Different definitions of preload are therefore:

- LVEDV (impossible to measure directly in the clinical setting)
- LVEDP (PCWP, or extrapolated even further away from the truth when using CVP and blood pressure). Here the LV **Compliance = δV/ δP** (change in volume/change in pressure) is important to keep in mind
- Ventricular end-diastolic wall stress – **diastolic La Place – T=Pr/h**
- EDPV point on the pressure volume loop

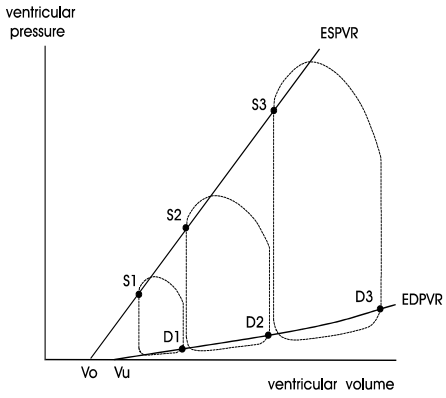


Figure: Preload and Afterload

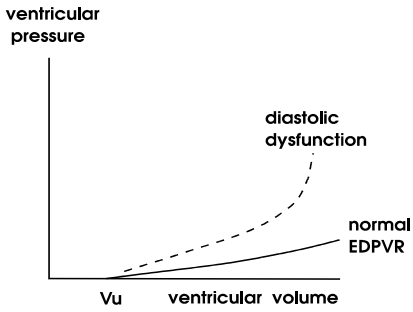


Figure: Compliance curve

:

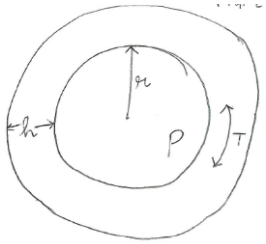


Figure: La Place

Afterload

In an isolated muscle fiber preparation, afterload is defined as the wall tension developed during contraction. Thus, afterload is related to the mechanical resistance to shortening of the muscle fiber. In the intact heart, afterload becomes the tension per unit cross section (T), developed in the ventricular wall during systole. This can be related to the intraventricular pressure during systole, by applying Laplace's law for pressure in an elastic sphere. Afterload is thus a measure of how forcefully the ventricle contracts during systole to eject blood.

The normal ventricle has an intrinsic ability to increase its performance in response to moderate increases in afterload, to maintain SV. If the afterload increases suddenly, it causes an initial fall in SV. The ventricle then increases its EDV in response to the change, which in turn restores the SV. This is called the Anrep effect.

Different definitions of afterload are therefore:

- Force resisting muscle fibre shortening
- Arterial impedance opposing LV ejection (Windkessel) - Systemic vascular impedance is the mechanical property of the vascular system opposing ejection and the flow of blood into it. This is composed of two components. One is the resistive or steady-flow component, which is the SVR. This component is mainly due to the frictional opposition to flow in the vessels. The other component is the reactive or frequency-dependent component, which is due to the compliance of the vessel walls and inertia of the ejected blood. This component is dependent on the pulsatile nature of the flow and rapidity of ejection. A major part of this reactive component is formed by the arterial elastance (Ea).
- Effective arterial elastance (Ea) - ventriculo-arterial coupling. Arterial elastance is the inverse of arterial compliance and is a measure of the elastic forces in the arterial system that tends to oppose the ejection of blood into it. Determination of Ea involves plotting a PV curve for the arterial system using different SV and recording end-systolic pressures. The slope of the curve then gives the effective elastance (compliance^{-1}) of the arterial system.
- Ventricular end-systolic wall stress – **systolic La Place – T=Pr/h**
- **Ohm's Law, R=P/Q** (Resistance=Pressure/Flow, dynes.sec.cm^{-5}) – using the PA catheter to calculate systemic vascular resistance. Systemic vascular resistance (SVR) is the most commonly used index of afterload in clinical practice, and can be calculated from mean arterial pressure (MAP), central venous pressure and CO, as follows:

  SVR=MAP - RAP/CO X 80 dynes.s.cm^{-5}

  The normal value for SVR ranges from 900 to 1400 dynes.s cm^{-5}. SVR is not a good estimate of afterload, since it is only one component determining afterload, and does not provide any index of intraventricular pressures generated during systole (i.e. how hard the ventricle is contracting). Clearly, if the ventricle only generates low intraventricular pressures by contracting softly, the afterload is low irrespective of the calculated SVR.

  In a similar manner the pulmonary vascular resistance (PVR) may be calculated as an index of RV afterload, using mean pulmonary arterial pressure (MPAP), pulmonary capillary wedge pressure and CO:

  PVR = MPAP – PCWP/CO X 80 dynes.s.cm^{-5}

  The normal PVR ranges from 90 to 150 dynes.s cm^{-5}. CO, PAP and PCWP have to be obtained with a PAC to calculate SVR and PVR.

3. Bernoulli Equation (and Doppler)

When a transvalvular velocity is measured with Doppler ultrasound, the Bernoulli equation is used to convert velocity into a pressure gradient (or pressure drop). To measure an accurate flow velocity with ultrasound, the rule of Doppler (Johan Christian Doppler, 1803-1853) states that the ultrasound beam of interrogation must be parallel to the blood flow, to avoid underestimation.

The average pressure drop/gradient to open a normal valve, i.e. aortic valve in systole, or mitral valve in diastole is anything between 2-4 mmHg.

This principle is well demonstrated with the Wiggers diagram:

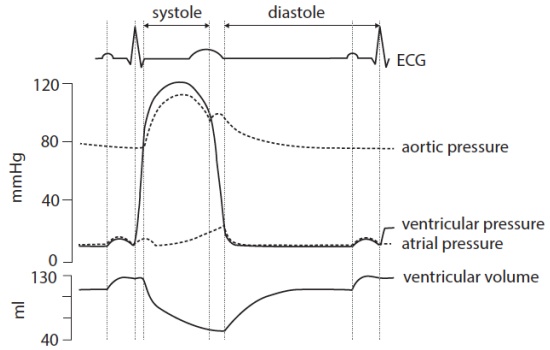


Figure: Wiggers diagram to explain pressure gradients across normal valves during systole and diastole

The Bernoulli equation in its full form has been described for flow dynamics (Daniel Bernoulli, 1738- "Hydrodynamica").

$$P_1 - P_2 = 0.5\rho\ (v_2^2 - v_1^2) + R(v) + r\int_1^2 dv/dt\ dS$$

Pressure Gradient	convective acceleration	viscous friction	flow acceleration

Equation: In its original form, the Bernoulli equation has three parts, which consider (i) convective acceleration where ρ (rho) refers to blood density, (ii) viscous friction and (iii) the rate of change of flow acceleration.

In the first part of the Bernoulli equation, blood density (ρ) multiplied by 0.5 is 3.98, but is rounded up to 4. The second and third parts of the original equation are then assumed to be more or less constant; hence the modified Bernoulli equation is created:

$$\textbf{Pressure Gradient} = 4 \times (V_2^2 - V_1^2)$$

Equation: The modified Bernoulli equation, where V_2 is the maximal velocity across the AV, and V_1 is the maximal velocity across the LVOT.

V_1 in the Bernoulli equation refers to the velocity upstream from the constriction (ie the velocity in the LVOT when being applied to AV velocities). If the LVOT velocity is less than 1 m/sec then the value for V_1 is assumed to be negligible and also ignored. This allows the simplified Bernoulli equation:

$$\textbf{Pressure Gradient} = P1\text{-}P2 = 4\ V^2$$

Equation: The simplified Bernoulli equation, where V is the maximal flow velocity across the AV.

If the velocity is higher than 1 m/sec in the LVOT (like e.g. in HOCM) then the value for V_1 cannot be ignored and the modified Bernoulli equation must be used.

The Bernoulli principle has many more assumptions than just those referred to in the equation above. For example, in aortic valve (AV) stenosis it is assumed that the diameter of the left ventricular outflow tract (LVOT) prior to the AV constriction is of the same diameter as that of the ascending aorta, and the surface area of the LVOT is the same as that of the ascending aorta. Other assumptions are that the inflow shape in aortic stenosis has a flat, orifice-like shape as opposed to a funnel shape, and that flow through the constriction is laminar, not turbulent.

Severe AS in a native valve is defined as a flow velocity more than 4.5 m/sec, a mean pressure gradient (mean pressure drop) of more than 50 mmHg and a peak pressure gradient (peak pressure drop) more than 80 mmHg. The mean pressure gradient is obtained by accurately tracing the outline of the velocity time integral signal, away from the transducer when Continuous Wave Doppler (CWD) is placed across the aortic valve. The mean gradient is more appropriate in reflecting the severity of pressure gradient as several factors can alter the peak velocity.

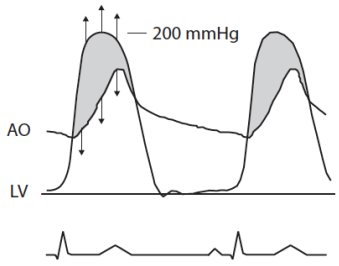


Figure: Pressure gradient during systole in Aortic Stenosis

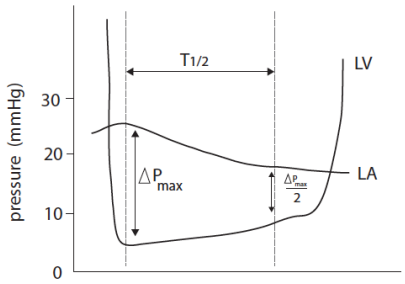


Figure: Pressure gradient during diastole in Mitral Stenosis

Compare the numbers of **severe AS** with that of across **normal valve** (PG = 4 x V^2) therefore more or less, flow velocity 1 m/sec, peak pressure gradient of 4 mmHg, mean pressure gradient of 2 mmHg.

4. Fick (Principle vs Law vs Equation), Shunt and Dead Space

The **Fick principle** states that the amount of a substance taken up by an organ (or the whole body) per unit time is equal to the arterial concentration of the substance minus the venous concentration (a–v difference), multiplied by the blood flow. This can be applied to the oxygen content of blood to determine Cardiac Output (CO).

First, the steady-state oxygen content of venous (CvO_2) and arterial blood (CaO_2) is measured. Then oxygen uptake in the lungs is measured over 1 minute (VO_2). Finally, the Fick principle is applied to calculate the blood flowing in 1 minute:

$CO = VO_2/(CaO_2 - CvO_2)$

Errors in sampling, and the inability to maintain steady-state conditions, limit this technique.

In clinical practice the intensivist applies the Fick principle when a pulmonary artery catheter (PAC) is used in the critically ill patient. It is important to distinguish between indices measured with the PA catheter (right atrial pressure, right ventricular pressure, pulmonary artery pressure, pulmonary capillary wedge pressure, cardiac output, etc.) and those indices derived and calculated from the measured parameters (e.g. stroke volume, stroke work index, and the vascular resistances). It would border on negligence to have a PA catheter in situ and ignore

the oxygen demand and supply parameters (mixed venous oxygen content, oxygen delivery, oxygen consumption, etc).

Rearranging the equation:

$CvO_2 = CaO_2 - VO_2/CO$

The saturation of the mixed venous blood (SvO_2) can therefore be measured in the RV or proximal PA, and give a good indication of peripheral oxygen consumption on cellular level, as well as the oxygen delivery ($DO_2 = CaO_2 X$ CO).

Over the years there have been innovative changes in design of the PA catheter using Fick principle, including special purpose PA catheters for monitoring continuous cardiac output, continuous mixed venous oximetry (SvO_2), etc.

The perfusionist also uses the Fick principle when caring for a patient on the cardiopulmonary bypass (CPB) machine during a heart operation. He/she monitors the drained from the right atrium (inline SvO_2), and then uses the Fick principle to balance the patient's oxygen supply and demand by adjusting machine blood flow (perfusion) rates.

Fick's Law of Diffusion

Fick was a clever physiologist-physicist who described not only a principle-equation, but also a law!

The transfer of a gas across the blood-gas barrier of the alveolar membrane is called diffusion. Diffusion across the alveolar membrane is well described by Fick's law, which states that the rate of transfer of a gas through a sheet of tissue is proportional to the tissue area (A) and the difference in gas partial pressures between the two sides (P1-P2), and inversely proportional to the tissue thickness (T). The area of the blood-gas barrier in the lung is enormous and the alveolar membrane is very thin, therefore ideal for diffusion of oxygen. The rate of gas transfer is also proportional to a diffusion constant (D), which involves the solubility of the gas (sol) and the square root of its molecular weight (mw).

$Vgas = A/T \times D \times (P1-P2)$ **$D = sol/\sqrt{mw}$**

The alveolar partial pressure of oxygen (PAO_2) is the driving force (P1) behind the diffusion of O_2 across the alveolar membrane into the pulmonary capillaries. In the capillaries the O_2 combines with haemoglobin in the red blood cells. When the haemoglobin is saturated with O_2, the partial pressure of oxygen in the pulmonary capillary blood (PcO_2) will rise. Under normal resting conditions the PcO_2 will almost reach that of the alveolar gas (PAO_2) when the red cell is about one-third of the way along the capillary. Therefore the O_2 uptake can be regarded as occurring in two stages:

- diffusion of O_2 through the blood-gas barrier including the plasma and red cell interior.
- reaction of O_2 with haemoglobin

The resistance to this reaction can be described by the following reaction:

$1/DL=1/DM + 1/q.Vc$

where DL is the diffusing capacity of the lung, 1/DM is the resistance of the blood-gas barrier, q is the rate of O_2 reaction with haemoglobin and Vc is the volume of pulmonary capillary blood.

In a perfect lung the PO_2 of arterial blood (PaO_2) would be the same as that in alveolar gas (PAO_2). In real life this is not true. As blood travels through the pulmonary capillary, its PO_2 rises closer and closer to that of alveolar gas. Under normal conditions there is always a small difference between alveolar and end-capillary PO_2 (A -a gradient), because of incomplete diffusion.

5. Hagen-Poiseuille

What is the Hagen-Poiseuille Equation?

The Hagen-Poiseuille Equation describes laminar flow through a tube:

$Q = Pr^4\pi/8\eta L$

Q = flow
P = pressure gradient across tube/vessel
r = radius of tube/vessel
η = fluid viscosity
L = tube/vessel length

How does laminar flow differ from turbulent flow? What is the Reynolds number (Re)?

With laminar flow the layers of fluid are moving smoothly, with the central "lamina" flowing fastest, and the outer layer slowest. Turbulent flow, on the other hand, is chaotic, with fluid eddies. The Reynolds number predicts whether flow will be turbulent or laminar. A Re < 2000 means laminar flow, while > 2000 predicts turbulent flow.

$Re = V\rho d/\eta$

V = fluid velocity
ρ = fluid density
d = vessel or tube diameter
η = fluid viscosity

What affects the rate of turbulent flow?

Flow is directly related to the square root of the pressure gradient ($\sqrt{P}$)
Flow is directly related to the square of the radius (r^2)
Flow is inversely related to the vessel or tube length (1/L)
Flow is inversely related to fluid density (1/ ρ)

The influence of the Hagen-Poiseuille principle is appreciated in daily clinical practice when discussing the respiratory airways, or considering the thickness and length of an endotracheal tube, ventilation/breathing circuits, fluid-giving sets, intravenous/arterial cannulas, Cardiopulmonary Bypass or ECMO Circuits, etc.

References

- Davis PD and Kenny GNC. Basic Physics and Measurement in Anaesthesia. 5th Edition. Oxford: Butterworth Heineman, 2003.
- Hauser N, Myburgh A, Swanevelder J. Chapter 2: Anatomy and Pathology of the heart and major vascular system. In Oxford Textbook of Cardiothoracic Anaesthesia, Editors Alston P, Myles P, Ranucci M. Oxford University Press 2015.
- Swanevelder JLC. Chapter 5: Cardiac Physiology, in Fundamentals of Anaesthesia, Pinnock C, Lin T, Smith T (editors). Greenwich Medical Media Ltd 2003
- Heart physiology and pathophysiology, 4th edition - Sperelakis N, Kurachi Y, Terzic A, Cohen M (editors). 2001 Academic press
- Cardiac Anesthesia, 4th edition - Kaplan JA (editor). 1999 WB Saunders
- Ventricular Function, SCA monograph - Warltier DC (editor).1995 Williams & Wilkens
- Review of Medical Physiology,18th edition - Ganong W.1997 Appleton & Lange
- Respiratory Physiology – the essentials. West JB. Ninth Edition.

Interpretation of acid-base abnormalities

Professor IA Joubert

Division of Critical Care
UCT Department of Anaesthesia & Perioperative Medicine
Ivan.Joubert@uct.ac.za

The following notes set out to provide essential information for use in the interpretation of blood gas printouts as encountered in the clinical environment. Basic concepts and definitions are emphasised, all of which are essential in the accurate interpretation of blood gasses.

The definition of pH

pH is defined as the concentration of hydrogen ions expressed as the negative log to base 10.

$$pH = -\log_{10}[H^+]$$

As an abnormality of pH indicates an abnormality of hydrogen ion concentration, interpretation of this variable alone, can only indicate either acidaemia or alkalaemia.

The blood gas analyser

From a pure blood gas perspective, a blood gas analyser only measures three variables – pH, pO_2 and pCO_2. For the purposes of the following discussion, only pH and pCO_2 are relevant.

All of the remaining variables are calculated – primarily through manipulation of the Henderson-Hasselbalch equation.

Important concepts for interpreting blood gasses

The Henderson-Hasselbalch equation

This equation expresses the relationship between pH, pK_a and [acid] and [base].

$$pH = pK_a + \log\frac{base}{acid}$$

For blood gas interpretation purposes this equation can be expressed as:

$$pH = pK_a + \log\frac{[HCO_3^-]}{[H_2CO_3]}$$

Or more simply:

$$pH = 6.1 + \log\frac{[HCO_3^-]}{0.03\ pCO_2}$$

Here 0.03 is the solubility coefficient of CO_2 when expressed in mmHg. [0.225 is the solubility coefficient expressed in kPa]. 6.1 is the pK_a of the bicarbonate buffer system.

Bicarbonate

As the blood gas analyser measures pH and pCO_2, the HCO_3 is a calculated variable.

Bicarbonate can only be used to assess metabolic disturbance in a patient with no respiratory abnormality!

As a calculated variable, bicarbonate is affected by both respiratory and metabolic disturbances. It cannot, therefore, be an ideal measure of either. Moreover, the relationship between metabolic acidosis and bicarbonate is neither consistent nor linear. And, finally, in acid-base determinations the concentration (in mEq/L) of the bicarbonate ion (HCO_3^-) is not measured, it is calculated from the pCO_2 and pH.

Bicarbonate is therefore not a particularly useful variable – it is merely the product of calculation!

Standard bicarbonate

While bicarbonate itself is a poor measurement of either the respiratory or metabolic regulator, **standard bicarbonate** is a better measurement of the metabolic component.

It was introduced in 1957 by Jorgensen and Astrup. It was defined as the bicarbonate concentration under standard conditions: pCO_2=40 mmHg (5.3kPa), temperature of 37^{o}C, and haemoglobin being fully saturated with oxygen.

As the standard bicarbonate includes correction for any respiratory abnormality, it is useful in the identification of metabolic disturbance.

Base excess

The year after introducing Standard Bicarbonate, Astrup and Siggard-Andersen in 1958 introduced Base Excess as a better method of measuring the metabolic component. In essence the method calculated the quantity of Acid or Alkali required to return the plasma **in-vitro** to a normal pH under standard conditions (these being pCO_2 and temperature).

Standard Base excess

Standard Base Excess is the Base Excess value calculated for anemic blood (Hb = 5 g/dl) on the principle that this closely represents the behavior of the whole human being. The rationale for this is that in the whole body, hemoglobin effectively buffers the plasma and the much larger extracellular fluid, i.e., the behavior is that of anemic blood. The method predicts the quantity of Acid or Alkali required to return the plasma **in-vivo** to a normal pH under standard conditions.

In the clinical arena, if standard base excess is available, it represents the **best** measure of the metabolic disturbance.

Assessing metabolic disturbance

From above it is appreciated that base excess and standard base excess can be used to identify metabolic disturbance. Remember that respiratory abnormalities are excluded in these calculations.

Nevertheless, neither of the above can explain the mechanism of any metabolic disturbance. Metabolic disturbance can only be explained through the calculation of anion gap, or better through the appreciation of strong ion difference and total ionized protein.

The Anion gap

This is calculated as:

$$Anion\ gap = [Na^+ + K^+] - [Cl^- + HCO_3^-]$$

The anion gap is only useful in the description of metabolic acidosis, where it is classified as either being increased, or decreased.

The causes of a raised anion gap acidosis are:

- lactic acidosis
- ketoacidosis
- chronic renal failure (accumulation of sulfates, phosphates, uric acid)

- intoxication or drug overdose, e.g., ethanol, methanol, ethylene glycol, formaldehyde, paraldehyde, salicylates, INH, toluene, sulfates, metformin.
- rhabdomyolysis

The causes of a normal anion gap acidosis (mostly associated with a Cl^- abnormality) are:

- Longstanding diarrhoea (bicarbonate loss)
- Uretero-sigmoidostomy
- Pancreatic fistula
- Renal Tubular Acidosis
- Intoxication, e.g., ammonium chloride, acetazolamide, bile acid sequestrants

Strong ion difference

The Strong Ion Difference is the difference between the sums of concentrations of the strong cations and strong anions:

$[SID] = [Na^+] + [K^+] + [Ca^{2+}] + [Mg^{2+}] - [Cl^-] - [\text{Other Strong Anions}]$

Lactate is a strong anion and should be considered in the above equation.

Strong ions are always completely dissociated in solution. Note that bicarbonate is not a strong ion at all!

The normal strong ion difference is about 35. Any departure from this number is roughly equivalent to the SBE.

Weak non-volatile acids - $[A_{TOT}^-]$

$[A_{TOT}]$ is the total plasma concentration of the weak non-volatile acids, inorganic phosphate, serum proteins, and albumin.

$[A_{TOT}] = [P_{iTOT}] + [Pr_{TOT}]$ + albumin.

Proteins provide a significant source of ionisable substrate that is useful in the buffering of acid-base disturbances. A low albumin plays an alkalinizing role from an acid-base perspective.

The use of the strong ion difference and abnormalities of A_{TOT}^- can provide additional insight into the appreciation of the cause of an acid-base disturbance.

Clinical considerations

Changes in acid-base status are either respiratory or non-respiratory, i.e., metabolic:

Respiratory:

The effects of changes of pCO_2 are well understood and produce the expected alterations in $[H^+]$:

$CO_2 + H_2O <—> H_2CO_3 <—> HCO_3^- + H^+$

Metabolic (Non-Respiratory):

Metabolic disturbances cannot be viewed as a consequence of bicarbonate concentration because bicarbonate is merely a dependent variable. The two possible sources of metabolic, i.e. non-respiratory, disturbances are either [SID] or $[A_{TOT}]$, or both.

With normal protein levels, [SID] is about 40mEq/L. Any departure from this normal value is roughly equivalent to the standard base excess (SBE), i.e., if the measured [SID] were 45 mEq/L, the BE would be about 5 mEq/L, and a measured [SID] of 32 mEq/L would approximate to a BE = -8 mEq/L. Because [SID] does not allow for haemoglobin, there is often a small discrepancy.

Changing [SID]:

[SID] can be changed by two principal methods:

1) Concentration:

Dehydration or over-hydration alters the concentration of the strong ions and therefore increases, or decreases, any difference. The body's normal state is on the alkaline side of neutral. Therefore, dehydration concentrates the alkalinity (contraction alkalosis) and increases [SID]; whereas, over-hydration dilutes this alkaline state towards neutral (dilutional acidosis) and decreases [SID].

2) Strong Ion Changes:

If the sodium concentration is normal, alterations in the concentration of other strong ions will affect [SID]:

a. Inorganic Acids:

The only strong ion capable of sufficient change is chloride, (potassium, calcium and magnesium do not change significantly). An increased Cl^- concentration causes an acidosis and a decreased [SID] – hyperchloraemic acidosis. Because the chloride ions are measured, the anion gap will be normal.

b. Organic Acids:

By contrast, if the body accumulates one of the organic acids, e.g., lactate, formate, keto-acids, then the metabolic acidosis is characterized by a normal chloride concentration and an abnormal anion gap because of the presence of the "unmeasured" organic acid.

Changing [A_{TOT}^-]:

The non-volatile weak acids comprise inorganic phosphate, albumin and other plasma proteins. Making the greatest contribution to acid-base balance are the proteins, particularly albumin, which behave collectively as a weak acid.

Hypoproteinemia, therefore, causes a base excess and vice versa.

Phosphate levels are normally so low that a significant fall is impossible. However, in renal failure, high phosphate levels contribute to the acidaemia.

Interpreting acid-base derangement

The initial inspection of pH, pCO_2 and SBE are likely to be most helpful.

Simple respiratory acidosis is easy to identify, most commonly resulting from a depression of minute ventilation, for a variety of reasons. The SBE is normal.

Respiratory alkalosis is relatively rare. Hyperventilation is an unusual physiologic disturbance that may be secondary to hypoxia (with no depression of minute ventilation), high altitude or unusual drive of the respiratory center. Once again the SBE is normal.

Interpreting metabolic disturbance is best done through the inspection of SBE, thus discounting any respiratory component of the disturbance to pH.

For metabolic acidosis further insight is gained through the use of anion gap, or SID and [A_{TOT}^-].

Primary metabolic alkalosis may occur for a variety of reasons:

- Loss of acid via: urine, stools, or vomiting
- Transfer of hydrogen ions into the cells
- Excessive bicarbonate administration, e.g. alkali given to patients with renal failure.
- Contraction of the extracellular space due to excessive diuretic treatment

Simple mathematics!

As a rule of thumb the following holds true:

pCO_2	pH	HCO_3^-
12 mmHg	0.1	6 mEq/l
1.6 kPa		

The equation means that a change of 0.1 in the pH can be caused by either:

1. A respiratory change (PCO_2 change) of 12mmHg, or
2. A metabolic change (Base Excess change) of 6 mEq/L.
3. A mixture of the two.

This relationship allows the components to be "added" and "subtracted". For example, a pH of 7.2 (0.2 more "Acid") can be caused by:

1. a PCO_2 of 64 with a BE = 0 mEq/L
2. a PCO_2 of 52 with a BE = -6 mEq/L
3. a PCO_2 of 40 with a BE = -12 mEq/L
4. a PCO_2 of 32 with a BE = -18 mEq/L

Although this relationship is an approximation, it provides acceptable clinical results in most circumstances; its real value is in granting insight and understanding.

Identifying compensation

Compensation for acid-base disturbances is never complete from a mathematical perspective. In other words the pH can never be brought back to 7.4 by physiologic means. Compensation may be complete in that physiology has done all it can to offset the disturbance. At best complete physiologic compensation will lie roughly halfway between full mathematical compensation, and no compensation.

Conclusion

Acid-base interpretation is easy!

- Identify an acidaemia or alkalaemia
    - If the pH falls within the normal range in the face of significant respiratory and metabolic disturbance consider that there may be added complexity! Remember that compensation is never mathematically complete.
- Then look at pCO_2 and SBE.
    - Start to identify either respiratory or metabolic causes of the disturbance
- For metabolic acidosis
    - Consider both the anion gap and SID for insight
- Consider SID and $[A_{TOT}^-]$ for all metabolic disturbances
- Use the rule of thumb relationship to help predict necessary compensation.
- For the last time – remember that Compensation is never mathematically complete – usually only half compensation is possible.

Respiratory Mechanics

Dr Ernest Welch

Private Practice
Honorary consultant- Wits University

Inspiration

Diaphragm moves downward and Intercostal muscles and ribs move upward and outward. This produces an increase in the vertical length and circumference of the chest cavity.

In normal respiration a Tidal volume (TV) of about 500ml is produced.
The diaphragm moves about 1cm downward producing about 100ml of this TV.
The circumference of the chest increases by about 1.3cm over its entire length (about 30cm) producing the remaining 400 ml needed to create a total TV of 500ml.

In active inspiration the diaphragm will only move down by a maximum of 3cm (about 300ml) while the circumference increase (10-12cm) is responsible for the remaining 3 -4000ml change in TV.

Exhalation is essentially passive but may be supplemented the abdominal muscles increasing the intra-abdominal pressure, forcing air out of the lungs. (The passive nature of exhalation is shown on a pressure volume curve for a lung) (See Below)

Measurements of lung mechanics

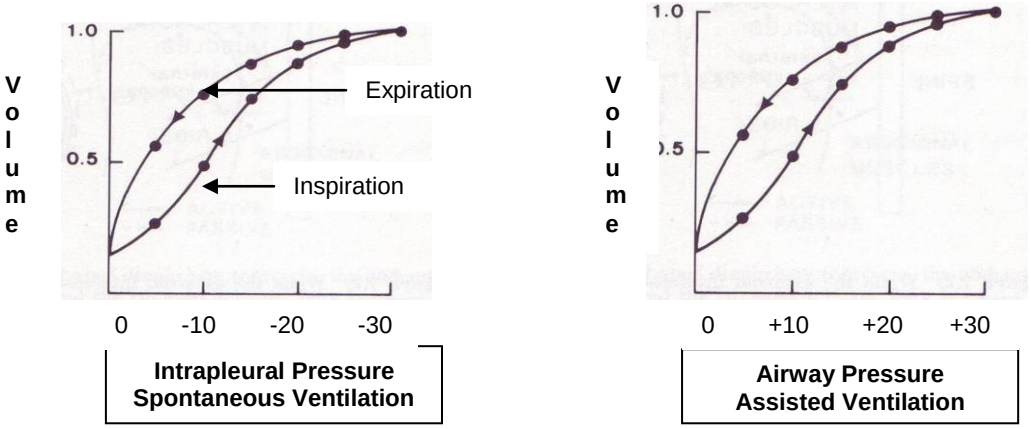
- **Static** - Measurements made when air is not moving (e.g. lung compliance)
- **Dynamic** - Measurements made determined when air is moving (e.g. FEV1)
- **Volume** – A measured number (e.g. Tidal volume)
- **Capacity** – A calculated number (e.g. Total lung Capacity = TV + RV +IRV + ERV)

Pressure volume relations of the lung

Spontaneous ventilation (SV) depends on the generation of negative pressure in the pleural space. As the chest wall moves away from the lung a space is produced which results in a negative pressure between the chest wall and the lung. The greater the distance the chest wall moves the more negative this pressure will be. This negative pressure results in air being sucked inwards into the lung.

In assisted ventilation (AV) air is forced into the lung pushing out the chest wall. An increase in pressure pushing the lung outward produces a bigger volume in the lung.

Therefore as the pressures inside the lung changes so does the volume.



Important things to notice on these curves

- There is no difference between the change in volume for SV and AV.
- The pressure in SV is negative and reflects the intrapleural pressure while AV is positive and reflects the pressure in the airways.
- In exhalation the curve follows a different path to inspiration – **Hysteresis** - due to air exiting the lung passively.
- There is always some air left in the lung after exhalation so inspiration never starts from zero (residual volume).
- Zero pressure is atmospheric pressure.

Compliance
Is the ability of the lung to accept volume as the pressure changes.
It is calculated as:

$$\frac{\text{Change in volume}}{\text{Change in pressure}} \qquad \text{(Normal is 200ml/cm } H_2O\text{)}$$

It is dependent on

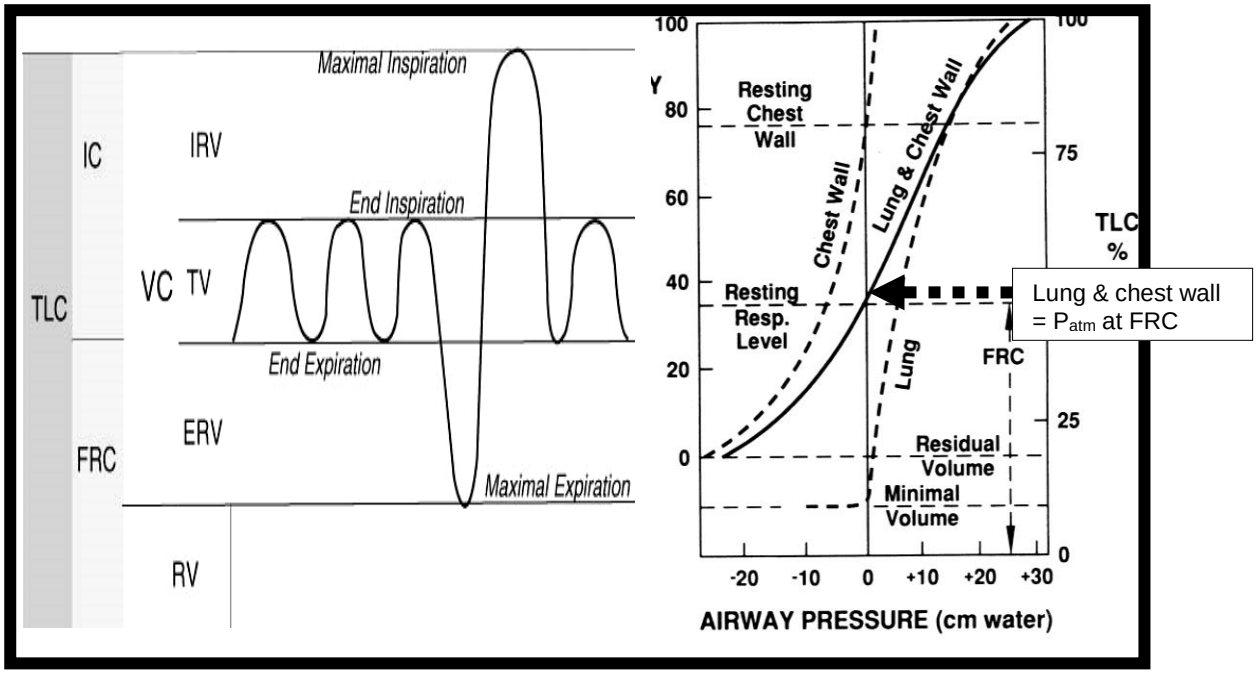
- The elastic stretch of the alveoli. (Static compliance). This is determined by collagen, elastin & air fluid interfaces (surfactant).
- The chest wall and pleural cavity rigidity. (Static compliance)
- The airway resistance to gas flow. (Dynamic compliance) Which is affected by the diameter of the conducting airway, and the work needed to open that airway.

Compliance is constantly changing, varying in different parts of the lung and with different pressures. It is also affected by oedema fluid in the alveolus and interstitium, atelectasis, loss of surfactant, fibrosis, age, COPD and chest wall disease.

Compliance and recall of the lung and chest wall in combination

Fundamentals
In a pneumothorax the lung collapses and the chest wall moves outwards!

1. The lung is always at a positive pressure (above P_{atm}) and wants to collapse.
2. The chest wall is always kept at a pressure less than P_{atm} and wants to spring outwards.
3. At FRC (functional residual capacity) the positive pressure in the lung balances the negative pressure in the chest wall and normal spontaneous respiration starts from here i.e. we breathe our tidal volume from atmospheric pressure.
4. With inspiration the chest wall is pulled outwards and it draws the lung out at the same time increasing the pressure in both
5. The chest wall pressure only reaches P_{atm} with deep active inspiration.



Airway resistance

Airflow
3 types of airflow are seen in the airways

1. Turbulent flow (In trachea)
2. Laminar flow (In very small airways)
3. Transitional (A mixture of Laminar and turbulent - in the majority of the respiratory tree)

Flow characteristics are determined by Reynolds number (Re) with radius and velocity of flow being the principal determinants.
(>2000 turbulent & <10 laminar)

$$Re = \frac{2rvd}{\eta}$$

r = radius
v = average velocity
d = density
η = viscosity
l = length
R = resistance
V = flow
P = pressure

From *Ohms* law we know that
Pressure (P) = Flow (V) x Resistance (R)
Or $R = \frac{P}{V}$

Poiseuille determined flow in a tube to be:

$$V = \frac{P \pi r^4}{8\eta l} \rightarrow \frac{V}{P} = \frac{\pi r^4}{8\eta l} \rightarrow \text{and} \quad \frac{V}{P} = \frac{1}{R}$$

therefore $R = \frac{8\eta l}{\pi r^4}$

Halving the radius increases resistance 16 times while doubling the length only doubles it.

Determinants of airway resistance (Inspiratory & expiratory)

1. Bronchial divisions
- Major site of resistance is the first 7 divisions.
- Small airways produce less than 20% of airway resistance

2. Lung volume
- Increased volume decreases resistance
- Additional volume opens alveoli and terminal bronchioles
- Small tidal volumes may produce increased closing capacity increasing resistance

3. Pulmonary vasodilation
- Distended blood vessels obstruct terminal bronchioles

4. Bronchial smooth muscle contraction
- Worsens resistance and may be reversed by bronchodilators

5. Density and viscosity of gases
- Helium-O_2 admixtures decrease density and resistance, improving flow

6. Dynamic airway compression
- Expiratory flow is effort independent at low flow rates due to terminal bronchial compression by intra-thoracic pressure.
- Terminal bronchi which are not cartilage supported collapse when external pressure reaches the pressure within the bronchus
- This is worsened by forced expiration, loss of elastic tissue in COPD, increased airway resistance and the equal pressure point approaching the alveolus as flow decreases.
- The inability of alveoli to empty increases the intra thoracic pressure further (stacking)

Work of breathing

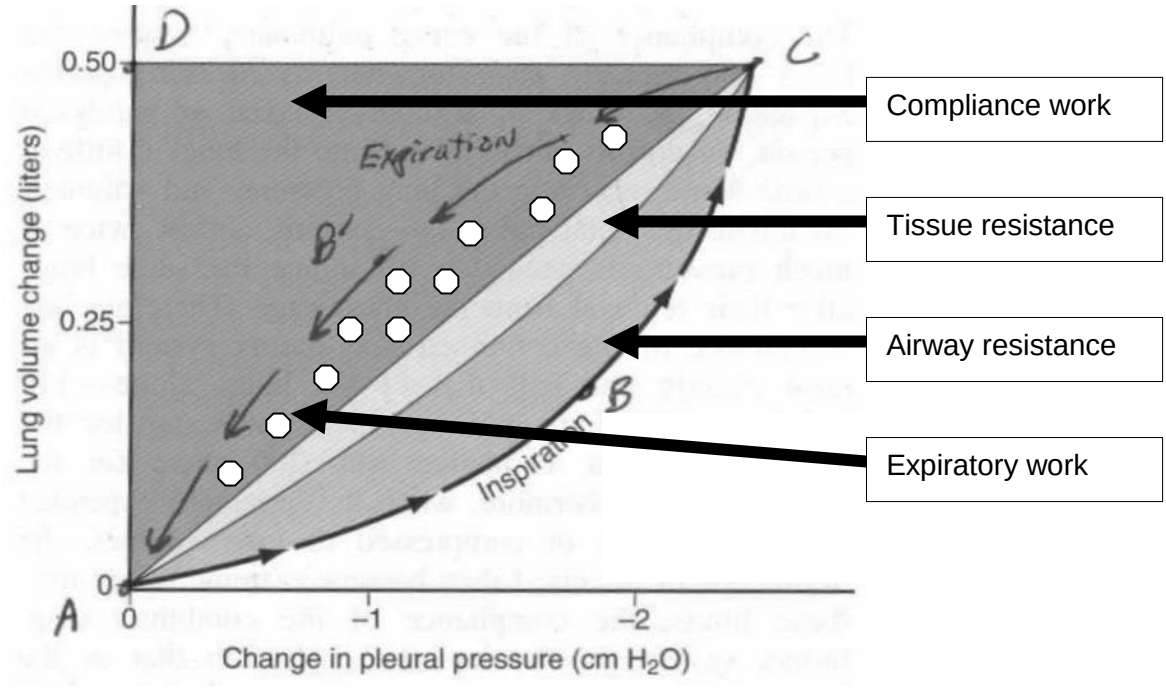
Work = force x distance
Force = pressure x area
Distance = volume/area
Work = (pressure x area) (volume/area)

Work = Pressure x Volume

The total work of is the sum of work needed to overcome:

- Compliance - essentially the work needed to overcome resistance from elastic forces of the lung and chest wall.
- Airway resistance
- Tissue resistance is the resistance provided by the lung pleura moving against the chest wall pleura. It is often added to airway resistance and called viscous resistance.

Work = Compliance + Airway resistance + Tissue resistance



Important things to notice from this curve:

1. Total work is all the areas added together
2. Exhalation work falls within the total work area and is done using the stored elastic energy from the compliance work
3. The compliance work less the expiratory work is dissipated as heat
4. Compliance work increases with higher tidal volumes
5. Viscous work (tissue & airway) increases with faster breathing and flow rates.
6. In COPD slow breaths with big tidal volumes takes advantage of increased compliance (less compliance work needed) while in restrictive airways disease fast and shallow breathing means less work against decreased compliance.

It is very difficult to measure work of breathing.
Work of breathing can be estimated from the percentage of O_2 consumption required for breathing.
Normal resting respiration requires less than 5%.
In COPD this can climb above 30% limiting effort tolerance.

References

- Respiratory physiology, John B West 6th edition
- Textbook of Medical Physiology, Guyton & Hall, 10th edition
- Anesthesia, Ronald D Miller 6th edition

Muscle Contraction

Dr Francois Roodt

UCT Dept of Anaesthesia & Perioperative Medicine

Introduction

Muscle – *contractile tissue*; Muscle are of 2 varieties: striated (striped) or nonstriated (smooth) muscle and make up to 50% of adult body mass.

Muscle tissue can be divided into 3 types of muscle:

i. skeletal muscle – attached to skeleton , functions as a unit,
morphologically striped muscle,
voluntary muscle under conscious control.
skeletal muscle fibre types:
- slow type muscle fibres
- twitch-type muscle fibres
  - fast fatigable fibres
  - slow fatigue-resistant fibres
  - fast fatigue-resistant fibres.

ii. cardiac muscle – myocardium, generate contraction of
atrial and ventricle muscle,
morphologically striped muscle.

iii. smooth muscle - involved in involuntary process (blood vessels, gut)
divided into unitary (visceral) and multi-unit smooth muscle (iris, ciliary muscle, piloerector muscles), morphologically nonstriped muscle.

Muscle properties[1-2]

Characteristics	Skeletal *m.*	Cardiac *m.*	Smooth *m.*
Anatomy:			
Cell shape/size	Long cylindrical cells, up to 30 cm long, 100 µm wide	Irregular, branched, rod-shaped cells, up to 100 µm long, 20 µm wide	Spindle shaped cells up to 400 µm long, 10 µm wide
Nuclei	Multinucleated	Mostly single	Single
Actin & myosin filaments	Yes	Yes	Yes
Sarcomeres present (striated)	Yes	Yes	No
Myogenic activity	No	Yes	Yes

- **sarcomere** is the basic unit of muscle contraction and is enclosed by adjacent Z-lines. Each sarcomere consists of one A band and two I band halves. Myosin forms the A band (thick filaments) and actin filaments (thin) form the I band. The actin filaments are anchored at the Z lines, with M lines formed by the connecting thick myosin filaments together. [3]
- **smooth muscle** – lack of cross striations, actin and myosin filaments arranged randomly.

A reductionist approach to skeletal, cardiac and smooth muscle morphology will be discussed during the presentation.

Characteristics	Skeletal *m.*	Cardiac *m.*	Smooth *m.*
Signaling:			
Initiation of contraction	Extrinsic (somatic, neural)	Intrinsic (muscle origin), but influenced by extrinsic autonomic (sympathetic & parasympathetic)	Intrinsic via nerve plexus, extrinsic via autonomic (sympathetic & parasympathetic), hormones or stretch
Depolarization/ *stimulation of contraction	Nerve action potential to each muscle cell, then to muscle endplate	Spread by specialized muscle cells in the conducting system, then spreads cell to cell	*Often spontaneous, sensitive to circulating chemical agents
Gap junctions (electronic coupling)	No, connective tissue separate adjacent cells	Yes (functional syncytium)	Yes – few in multi-unit, many in unitary
Hormonal influence on contraction	Small	Large	Large
Effect of nerve stimulation	Excitatory	Excitatory or inhibitory	Excitatory or inhibitory
Extent of innervation	Each cell innervated	Variable	Almost every cell in multi-unit, sparse in unitary
Spontaneous electrical activity	No	Yes	No in Multi-unit, Yes in unitary
Resting membrane potential	Stable RMP	Stable RMP in ventricular muscle. Not stable RMP in SA and AV node due to automaticity/rhythmicity	No RMP as potential tends to wander
Control	Voluntary control	Involuntary (and automatic contractions due to pacemaker cells)	Involuntary control only

- **gap junction** – also called a electrical synapse and provide connections between the cytoplasm of adjacent cells.

  It consists of a functional unit called connexion.

  This electrical conduction is much more rapid than the chemical synapses at the NMJ as it allows for:
  a) rapid depolarisation from cell to cell
  b) passage of ions and small molecules between cells.

- **motor unit** – consists of a single anterior horn α-motor neurone, its axon and all the muscle fibres it innervates, it is considered the functional unit of contraction.[3]

- **cardiac muscle has several special intrinsic properties –** automaticity, rhythmicity, conductivity, contractility and excitability. The nerves innervating the heart can only speed up or slow down the rhythm (chronotropy) and can modify the force of contraction (inotropy).

- **smooth muscle** – more sensitive to circulating chemical mediators.

Characteristics	Skeletal *m.*	Cardiac *m.*	Smooth *m.*
Muscle contraction physiology:			
Major source of Ca^{2+}	Sarcoplasmic reticulum	ECF and sarcoplasmic reticulum	ECF (sarcoplasmic reticulum poorly developed)
Ca^{2+} binds to	Troponin C	Troponin C	Calmodulin
Function of this binding by Ca^{2+}	Removes the inhibition of troponin I and exposes myosin binding sites	Removes the inhibition of troponin I and exposes myosin binding sites	Ca^{2+}-calmodulin activates myosin light chain kinase
Mechanism of excitation-contraction coupling	Via action potentials and t-system	Via action potentials and t-system ('calcium-triggered calcium release')	Via action potentials, calcium channels and/ or 2nd messengers
Activation of myosin ATPase	Phosphorylation not required	Phosphorylation not required	Requires phosphorylation which is catalyzed by active myosin light chain kinase
Duration of muscle contraction	Brief – 7,5 – 100 msec. depending on fibre type	300 msec. in the ventricle at normal heart rate	Last for long periods (tonic contraction)
Type of contraction	Phasic	Rhythmic	Tonic with some phasic
Basic muscle tone	Neural activity	None	Intrinsic and extrinsic factors
Speed of contraction	Fast	Slow	Very slow
Muscle spindles	Present and important in regulation of contraction	Not present	-

- Excitation-contraction coupling collectively refers to the sequential series of steps from the action potential to the subsequent muscle contraction.

Molecular machinery

Muscle contraction requires interaction between Ca^{2+}, troponin, tropomyosin, actin and myosin.

Myosin is a large complex protein, consisting of a long tail and 2 globular heads which each bind actin and ATP.

Two F-actin chains twisted together make up the thin filaments. Actin potentiates the ATPase activity of myosin.

Tropomyosin prevents interaction of myosin with actin and is modulated by troponin.
Troponin is a globular protein present on the thin filament (1:7 actin molecules), with 3 sub-units Troponin-T (binds tropomyosin), troponin-I (inhibits actomyosin ATPase) and troponin-C (binds Ca^{2+}).

Ca^{2+} alters the troponin-tropomyosin complex configuration, so that myosin can interact with actin. Ca^{2+} binds to troponin C on the thin actin filaments. This interacts with troponin I and tropomyosin and this complex moves away to expose the sites on actin where myosin heads bind. Actin and myosin form a cross-link.

Both muscle contraction and relaxation are energy processes requiring ATP.

Muscle contraction persists until the cytoplasmic Ca^{2+} concentrations fall, Ca^{2+} is released from troponin C and actin is unable to interact with the myosin heads. This occurs due to rapid pumping of Ca^{2+} back into the SR via Ca^{2+}-Mg^{2+} ATPase pump, and Ca^{2+} diffuses to the terminal cisternae (t-tubules), ready for release with the next contraction.

Mechanism responsible for skeletal muscle contraction

This can be explained by the sliding filament theory. Muscle contraction occurs because of the sliding of the thin filaments and thick filaments along each other. This sliding is produced by myosin head cross-bridges pulling the actin fibers toward the centre of the sarcomere. The sliding is powered by the hydrolysis of ATP by the ATPase activity of the myosin heads.

The sliding filament theory is confirmed by the observation that developed tension during isometric contraction depends on initial muscle length.

Skeletal muscles are arranged near their optimal length.

Cardiac action potential and contraction

In the absence of extracellular Ca^{2+}, cardiac muscle doesn't contract (Ca^{2+} triggered Ca^{2+} release). Cardiac muscle cannot exhibit tetanic contraction due to a prolonged refractory period.

Smooth muscle contraction considerations

There is twice as much actin and tropomyosin in smooth muscle compared to striated muscle, but no troponin.

- **calmodulin** – an intracellular Ca^{2+} binding protein, with 4 Ca^{2+} binding sites. Calcuim binding to calmodulin can activate 5 different calmodulin-dependant kinaseses – myosin light chain kinase, phosphorylase kinase, Ca^{2+}-calmodulin kinase I,II, and III.
  Smooth muscle contraction is initiated by this activation of myosin light chain kinase, resulting in phosphorylation of myosin and causes the cross-bridge mechanism to operate.
- **latch bridge mechanism** – a sustained smooth muscle contraction, where the myosin heads remain attached to actin despite becoming dephosphorylated and the fall in cytoplasmic Ca^{2+} concentrations.

Smooth muscle tends to contract in response to increased tension.

- **plasticity** – refers to the variable tension that develops in visceral smooth muscle at a given length.

Nitric oxide – plays an important role in vascular smooth muscle; (i) local control of vascular tone and (ii) contributes to vascular wall thrombo-resistance. In the latter it inhibits platelet aggregation and adhesion to the vascular wall.
It is produced from L-arginine in the vascular endothelium by nitric oxide synthase, and diffuses to the vascular smooth muscle. It stimulates guanylate cyclase forming cGMP, causing relaxation of vascular smooth muscle.

Topics for interactive discussion during the 2nd part of the presentation and/or further consideration:

i) What is a muscle spindle?
ii) Describe the stretch reflex?
iii) Physiology of neuromuscular transmission, action potentials and developed tension in skeletal and smooth muscle.
iv) Drugs affecting excitation-contraction coupling in cardiac and skeletal muscle.

References

1. Ward JPT, Physiology at a glance, 2005, Blackwell Publishing Ltd.
2. Power I, Principles of physiology for the anaesthetist, 2001, Arnold publishers.
3. Brandis K, The physiology viva Questions & Answers.
4. Yentis SM, Anaesthesia and intensive care A-Z, 2005, Butterworth-Heinemann.
5. Pinnell J, et al. Cardiac muscle physiology. Contin Educ Anaesth Crit Care Pain 2007; 7 (3):85-88.
6. Hopkins PM. Skeletal muscle physiology. Contin Educ Anaesth Crit Care Pain 2006; 6(1):1-6.

Neuromuscular blocking agents and their reversal

Dr Ernest Welch

Private Practice
Honorary consultant- Wits University

The introduction of neuromuscular blocking agents (NMBAs) produced vastly superior operating conditions, facilitating tracheal intubation and maintaining muscle relaxation during surgery. The down side of this improvement in surgical conditions is the anaesthesia nightmare of awareness and inability to ventilate and intubate a paralysed patient. It is also a challenge to completely eliminate the effect of the NMBA at the end of surgery and residual neuromuscular blockade may lead to post-operative complications. These agents are associated with an increase in the risk of mortality, due principally to the need to manage the airway, their unpredictable nature and often poor post-operative recovery facilities where residual blockade and its sequelae may go undetected.

The recovery from neuromuscular blockade (NMB) remains markedly variable between patients, due to individual patient pharmacokinetics, adjuvant agents administered, varying drug metabolism and timing of the drug dose. As a result, many patients experience residual NMB during the early post-operative period, which can result in significant patient morbidity. Therefore, a more certain method for muscle relaxation during surgery is required. The ideal muscle relaxant is an agent that not only has a rapid onset, but a controllable duration of action, ranging from a few minutes to hours which can be quickly terminated, while having no side effects. No currently available agent or combination of agents meets these criteria.

Succinylcholine is still in use, as it probably comes closest to fulfilling these criteria, with its rapid onset and short duration of action. However, its side-effect profile and risk of prolonged blockade in patients with deficiencies in butyrylcholinesterase (pseudocholinesterase) have led to suggestions that it should be withdrawn from routine use in anaesthesia.

The non-depolarising muscle relaxants currently in use have a much slower onset and longer, very unpredictable offset, but they do have the advantage of a much more favourable safety profile than succinylcholine. Thus, NMB is currently, usually achieved using NMBAs that require administration of a reversal agent to improve the predictability of offset at the end of surgery. However, NMB reversal is not without problems of its own: Reversal utilizes a competitive mechanism, using acetylcholinesterase inhibitors, such as neostigmine. This process allows the NMBA to be freed from the neuromuscular cleft by increasing the concentration of acetylcholine at the neuromuscular junction, promoting metabolism of the NMBA by the liver, kidneys or plasma esterases. The result is both unpredictable termination of the NMB and an excess of acetylcholine which not only affects the nicotinic receptors, but also muscarinic receptors causing a wide range of side effects, such as bradycardia, nausea, vomiting, abdominal cramps, excessive secretions, bronchospasm and miosis. These in turn must be treated with the concomitant administration of glycopyrrolate or atropine.

Timing of reversal administration is important as the onset of action of acetylcholinesterase inhibitors is slow and a certain degree of spontaneous recovery is required for these agents to be effective. In addition, profound NMB may not be reversible through this competitive mechanism.

Sugammadex is a novel reversal agent which has recently been introduced onto the South African market. It directly binds rocuronium, vecuronium and, to a lesser extent, pancuronium. This agent has the potential to improve the predictability of NMBAs.

Classification

1. Depolarising muscle relaxants
    - Quaternary amines – Suxamethonium

2. Non depolarising muscle relaxants
    - Amino steroids (Pancuronium, vecuronium, rocuronium)
    - Benzylisoquinolium diesters (Atracurium, cisatracurium, mivacurium,)

Mechanism of action

- **Depolarising muscle relaxants**

Succinylcholine has a double methonium group that resemble the structure of acetylcholine. These methonium heads bind acetylcholine receptor α subunits mimicking the action of acetylcholine to cause depolarisation of the muscle. These subunits remain bound, preventing further contraction until the succinylcholine molecule is metabolised.

- **Non depolarising muscle relaxants**

These massive agents bind a single site on the acetylcholine receptor and prevent further binding to the receptor due to their size.

Potency and onset

A sufficient number of post synaptic receptors (75%) must be blocked to produce neuromuscular paralysis. If an agent is particularly potent this will occur slowly as the amount of agent needed to produce a profound block is low and the result is that only a few receptors will be blocked per unit time resulting in a slow onset. Due to the agents potency if a quicker block is required, more agent must be given increasing the block duration.
Pancuronium and cisatracurium are the most potent agents available to us (Intubation dose =0.1mg/kg) and have onset times of around 4 minutes while rocuronium (Intubation dose = 0.6mg/kg) has an onset of less than 90 seconds.

Offset

Offset is initially dependent on redistribution, but after repeated ("top up") or large doses it is totally dependent on elimination, which can be aided to some degree by reversal. In general offset follows the following pattern:

- Ultra-short acting agents (less than 15 minutes) – Plasma esterase metabolised (Succinylcholine and mivacurium)
- Short acting agents (15 to 30 minutes) – Hofmann elimination and plasma esterases (Atracurium)
- Intermediate acting agents 30 to 60 minutes) - Liver metabolised (Vecuronium (40%), rocuronium (70%))
- Long acting agents (60 minutes plus) – Renal excreted (Alcuronium and pancuronium (80%))

Hofmann elimination was first described in 1841 in which carbon radicals were eliminated from quaternary molecules at high temperature and pH. In atracurium and its more potent isomer cis-atracurium an acyl group breaks away from the molecule at physiological pH and temperature leaving laudanosine, a compound with potential of neurotoxicity in large amounts.

Neuromuscular blocking drugs due to their long duration of action and rapid accumulation are not ideal agents to be used in infusions. The ability to comprehensively reverse blockade with sugammadex may introduce this as an alternative technique where complete, prolonged paralysis is desired.

Summary of current NMBAs clinical properties

Drug	2 x ED_{95} mg/kg	Onset (min)	Offset (25% recovery) min
Quaternary amines			
Succinylcholine	1.0	45sec	7.6
Benzylisoquinolium diesters			
Atracurium	0.4	2.5	38
Mivacurium	0.15	1.8	16
Cis-atracurium	0.1	7.7	46
Amino steroids			
Pancuronium	0.1	2.9	86
Vecuronium	0.1	2.4	44
Rocuronium	0.6	1.0	43

Doses

ED_{95} – is the dose needed to produce 95% twitch depression. As a safety measure to ensure we have optimal intubating conditions 2x ED_{95} is used as the intubating dose. Additional "top up doses" of 10% of the initial dose are usually used producing around 10 minutes of additional paralysis.
Dosing alterations should be considered in the following conditions:

- Females appear to be more sensitive to the effects of neuromuscular blockade with aminosteroidal compounds.
- Paediatric responses to neuromuscular blocking drugs differ between age groups:
    - Neonates and infants less than 2 years require adult doses, while older children requires doses up to 2 times greater than adults with top up doses given more frequently.
- Elderly
    - Initial dose is unchanged, with no change in onset time. Duration of action is longer because of slowed metabolism. Top up doses should be decreased and given less frequently.
- Obesity: Dosing is calculated using ideal body weight.

Indications for use:

1. Intubation, especially in rapid sequence techniques.
2. To provide surgical paralysis.
3. To ensure a patient does not move during precision surgical techniques.
4. To ventilate a patient using non physiological ventilatory modes such as prone positioning or high frequency jet ventilation.

Neuromuscular monitoring

Monitoring of neuromuscular blockade is often neglected but has not been shown to decrease the incidence of post-operative paralysis, but is the only method of detecting it.
Monitoring is recommended when using sugammadex, to determine the sugammadex dose.

How to select a neuromuscular blocking drug:

In choosing which drug to use in a particular case the following 4 questions need to be matched to the individual agent's characteristics.

1. How fast do I need to get this patient paralysed?
2. How long will the paralysis need to last?
3. How is the drug metabolised and will this patient be able to metabolise this drug in that way?
4. Is there any side effect that may harm this patient, or unique property that may benefit them?

Non-depolarising muscle relaxants

Atracurium (Tracrium®)

Atracurium, an intermediate acting drug with organ independent metabolism is broken down by Hofmann elimination and plasma esterases, at physiological pH and temperature. It is made up of 10 isomers. It does not display cumulative effects with repeated doses. It can cause histamine release especially in larger doses, resulting in hypotension, bradycardia and bronchospasm while allergic reactions, skin rash and hypertension have been reported.

Cisatracurium (Nimbex®)

The cis-cis isomer of atracurium, which has many of the same properties as atracurium, despite being 3 times more potent. Its slow onset and long duration of action, coupled with its organ independent elimination, make it an ideal drug for long term use in ICU. Cisatracurium has similar elimination to atracurium. Like atracurium, cisatracurium is reversed by neostigmine but not sugammadex.

Mivacurium (Mivacron®)

Mivacurium is a short acting drug with a long onset time, properties that have hampered its clinical usefulness. Metabolism is similar to succinylcholine through pseudocholinesterase (butyrylcholinesterase) in plasma. It is therefore also contraindicated in "scoline apnoea". It lends itself to administration via infusion due to its short half-life and lack of accumulation. Doses of 8 -10 µg/kg/hr

will maintain moderate paralysis. Neostigmine has been shown to accelerate recovery from a block with mivacurium but recovery is so rapid it is rarely used. Sugammadex has no effect on mivacurium. Histamine release may be seen after rapid administration of mivacurium.

Vecuronium (Norcuron®, Muscuron®)

Vecuronium is an intermediate acting aminosteroid having a similar structure to pancuronium but needs to be diluted in sterile water prior to administration. With repeat doses a small cumulative effect is exhibited. About 40% of vecuronium is metabolised in the liver with the remainder being excreted unchanged via the kidney and bile.
Vecuronium is reversed by both neostigmine and sugammadex. Sugammadex has a slightly lower affinity for vecuronium than rocuronium (90% vs 95%).

Rocuronium (Esmeron®)

Rocuronium is a steroidal agent with a fast onset and intermediate duration of action. It is a versatile NMBA having dose dependent onset times ranging from 3 minutes with a 0.3mg/kg dose to 45 seconds with a 1.2mg/kg dose. It is used as an alternative to suxamethonium in rapid sequence inductions where its action can be terminated by reversal with 16mg/kg of sugammadex almost immediately. Recovery to all end points for all doses displays wide variation. Accumulation occurs with both repeat bolus dosing and infusions. Repeat maintenance doses are 0.1mg/kg at around 20 minute intervals. It can be reversed by both neostigmine and sugammadex with sugammadex being designed specifically to bind and reverse rocuronium.

Tachycardias have been described with doses of 1.2mg/kg and above and allergy and anaphylaxis has been reported with a preponderance in certain nations (France, Australia and Norway).

Pancuronium (Pavulon®)

Pancuronium was a mainstay of muscle relaxation from its introduction in 1967 to the early part of the 21st century. It is still used for cases requiring prolonged muscle relaxation or where tachycardia may be beneficial. It is slow acting with a long duration of action due to renal excretion (80%). It can be reversed by both neostigmine and sugammadex. Sugammadex only binds 60% of pancuronium. Like all drugs having long half –lives, reversal should only be attempted once evidence of spontaneous recovery is present to prevent recurarisation.

Alcuronium (Alloferin®)

A long acting still used sporadically around the world with a very slow onset and offset due to its renal excretion. It accumulates with repeat doses and releases histamine causing bronchospasm, hypotension and a reflex tachycardia.

Depolarising muscle relaxants

Suxamethonium (Succinylcholine) (Scoline®)

It is the only depolarising agent available. It has the shortest onset and duration of action of all neuromuscular blocking agents, and organ independent metabolism. It produces unmatched intubating conditions and the ability to breathe soon after an unsuccessful attempt. It can be given intravenously, intra-muscularly and as an infusion. If it were not for its potentially fatal side effect profile it would be the perfect muscle relaxant. (Hyperkalaemia, trigger of malignant hyperthermia, Scoline apnoea, muscle pains, bradycardia, allergy and anaphylaxis.)

Reversal of neuromuscular blockade

The current method of reversal of neuromuscular blockade involves the use of indirect acting competitive acetylcholinesterase (AChE) inhibitors that raise the level of ACh in the synaptic cleft.
ACh competes with neuromuscular blocking agents in the neuromuscular junction.
The increase in ACh displaces the neuromuscular blocking agent, which moves into blood from where it can be metabolised and eliminated by plasma esterases, the liver and kidneys.
AChE inhibitors do not metabolise or inhibit neuromuscular blocking agents.

Neostigmine

Neostigmine is the only acetylcholinesterase inhibitor available. It acts in an indirect competitive manner. Neostigmine binds acetylcholine esterase preventing the metabolism of acetylcholine which then competes with the neuromuscular blocking agent to displace it from the neuromuscular junction allowing it to be metabolised in the blood, liver or excreted by the kidneys. A dose of 0.04 - 0.07mg/kg is effective in about 1 minute with a peak effect at 9 minutes. It is given with Atropine (0.1mg per 1mg neostigmine) or glycopyrrolate (0.2mg per 1mg neostigmine) to counteract the excess Ach producing muscarinic side effects. Glycopyrrolate is preferred as it has a slower onset of action, producing less tachycardia, less central nervous system effects with a longer duration of action.

Due to its competitive action neostigmine should not be administered until spontaneous recovery to at least 2 twitches on a TOF has occurred.

Sugammadex (Bridion®)

Is a selective relaxant binding agent designed to bind rocuronium (95%) irreversibly, and due to their similar structures vecuronium (90%) and pancuronium (61%) are also encapsulated, but in lesser amounts, resulting in slightly slower recovery. Sugammadex and its complex are not metabolized and are eliminated exclusively by renal excretion or dialysis in renal failure. Atracurium, its isomers, suxamethonium and mivacurium, are hardly bound by sugammadex.

Dosing
3 dose regimens are recommended, depending on the degree of recovery of blockade.

- Moderate block - 2mg/kg if at least 2 twitches are present on a TOF after rocuronium or vecuronium. Recovery to TOF of 0.9 occurs in about 2 minutes.\
- Deep or profound blockade - 4mg/kg with 1 or 2 twitches on a PTC for rocuronium or vecuronium. Recovery to TOF of 0.9 occurs in about 3 minutes.
- Immediate or emergency reversal of rocuronium - 16mg/kg. Recovery to a TOF ratio of 0.9 takes about 90 seconds when sugammadex is given 3 minutes after 1.2mg/kg of rocuronium.

Administration of neuromuscular blocking agent after sugammadex:
If a patient needs to be re-paralysed after a dose of sugammadex, it is recommended that rocuronium, vecuronium and pancuronium are not used within 24 hours. Atracurium and its isomers will be effective. Administration of rocuronium (1.2mg/kg) 5 minutes after reversal with 4mg/kg sugammadex is effective, but results in a slower onset [mean 3 minutes (2-5 minutes)] and shorter duration [mean 25 minutes (17-46 minutes)] of blockade.

Its *side effect profile* appears to be relatively benign: Coughing, bucking and movement while under anaesthesia have been reported due to the sudden reversal of block. Hormonal contraception is bound by sugammadex and should be treated like a missed dose of the oral contraceptive pill. Its clinical effect on coagulation and its allergenic potential have not yet been fully described. Isolated cases of its use in aborting rocuronium-induced anaphylaxis have been described.

Further reading

1. Applied pharmacology in Anaesthesiology and critical Care, Welch EH, Neuromuscular blocking agents 485-519. Editors Milner A and Welch E. Medpharm 2012
2. Lee C. Goodbye suxamethonium! Anaesthesia 2009; 64: 73-81
3. Anesthesia 7th edition 2010. Naguib M, Lien CA. Pharmacology of muscle relaxants and their antagonists. Editor Miller RD, Churchill Livingstone Elsevier.
4. Srivastava A, Hunter JM. Reversal of neuromuscular block. BJA 2009; 103: 115-29

Anaphylaxis
Clinical presentation and perioperative management

Dr Adri Marais

UCT Dept of Anaesthesia & Perioperative Medicine

Definition: "Anaphylaxis is a serious allergic reaction that is rapid in onset and may cause death"

(Summary Report-Second National Institute of Allergy and Infectious Disease/Food Allergy and Anaphylaxis network symposium)

The European Academy of Allergology and Clinical Immunology committee recommended that the term *anaphylactoid*, introduced for non–IgE-mediated anaphylactic reactions, should no longer be used (not universally accepted yet).

Introduction

Anaphylaxis during anaesthesia is a rare phenomenon, but may have life threatening consequences when encountered and if not managed appropriately. The anaesthetist is usually alerted to a crisis only when it is severe enough to cause rapid cardiovascular and respiratory compromise.

Early signs and mild symptoms remain virtually unrecognised when patients are unconscious and covered with surgical drapes, preventing observation of the initial skin manifestations. Secondly, the severity of the reaction may be underestimated by the anaesthetist. The cardiovascular deterioration may initially be masked by a light plane of general anaesthesia (or an extensive regional block). Conversely, hypotension and difficulty in ventilation may have other more common causes that need to be excluded first.

Multiple drugs are administered over a short period of time and allergenic agents are not limited to intravenous drugs or fluids. It may include other substances used in the operating room such as skin disinfectants, latex gloves and catheters. Skin or mucosal application leads to a delayed onset of reaction, often presenting 15-30 minutes into a procedure.

Pathophysiology

Anaphylaxis is an immediate immunologically mediated allergic reaction to an administered substance. It is classified as a type 1 hypersensitivity reaction involving multiple organ systems. It can be IgE- or non IgE-mediated.

The clinical manifestations are due to the immediate as well as on-going release of preformed mediators from mast cells and basophils. Initial sensitisation occurs when T lymphocytes in susceptible patients are presented with an allergen and in response produce **IgE antibodies**. The IgE antibodies bind to high affinity FcεRI receptors on mast cells and basophils (as well as low affinity FcεRII of leucocytes, platelets and eosinophils). This initial phase of sensitisation is clinically silent.

Re-exposure to the allergen cause multimeric cross-linking of the IgE antibodies, activating intracellular transduction cascades with the release of preformed mediators (histamine, tryptase, chymase and heparin) from mast cells and basophils. This triggers the release of pro-inflammatory phospholipid derived mediators (prostaglandin D2, leukotrienes, platelet activating factor (PAF), thromboxane A2). These mediators in turn cause the release of chemokines and cytokines, with the recruitment of inflammatory cells. A very small amount of antigen is needed to activate this severe allergic cycle.

The involved target organs are usually skin, mucous membranes, cardiovascular and respiratory systems, as well as the gastrointestinal tract. The corresponding clinical signs are described by the Ring and Messmer clinical severity scale (Grades I – IV).

Grade	Symptoms			
	Skin	*GI*	*Respiratory*	*Cardiovascular*
I	Local pruritis Flushing Urticaria Angioedema			
II	Same as above	Nausea Cramping	Rhinorrhea Hoarseness	Tachycardia (>20 bpm) Blood pressure Δ (>20 mmHg systolic) Arrhythmia
III	Same as above	Vomiting Defecation Diarrhea	Laryngeal edema Bronchospasm Cyanosis	Shock
IV	Same as above	Same as above	Respiratory arrest	Cardiac arrest

Non IgE-mediated immunologic type 1 reactions are clinically indistinguishable from IgE-mediated reactions. It can however occur on first exposure to the allergen. **IgG-mediated** reactions are much less frequent and less serious than IgE-mediated reactions. Histamine release may also be idiopathic or triggered directly (physical factors like cold or heat, drugs like morphine, atracurium, meperidine and vancomycin), or may be released in response to bradykinin or complement activation. The clinical response depends on both the drug dose and the rate of delivery, but is usually benign and confined to the skin.

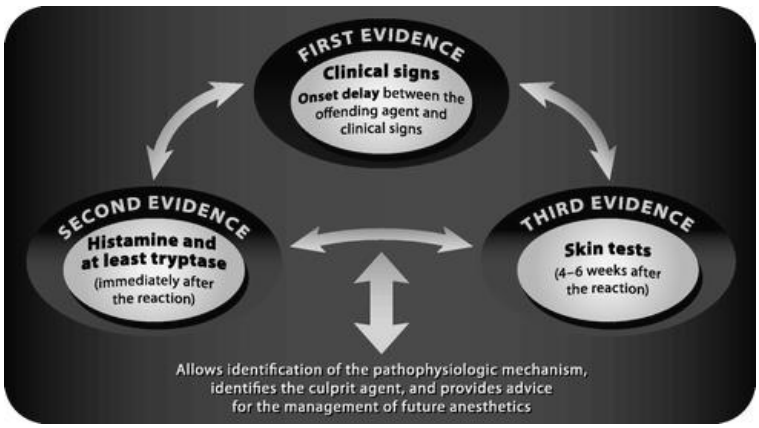
Epidemiology

The overall incidence of perioperative anaphylaxis is estimated at 1 in 10 000 – 20 000 anaesthetic procedures. It is estimated at 1 in 6500 administrations of neuromuscular blocking agents. Females are more affected than males. The exact incidence remains underestimated as reactions are underreported. Morbidity rates due to perioperative anaphylaxis remain unknown, but between 3 -10% of mortality rates (France, United Kingdom) that are partially or totally anaesthesia related involves anaphylaxis.

Allergic reactions to every drug used in anaesthesia (except the volatiles) have been documented, with neuromuscular blocking agents (NMBAs) followed by antibiotics and latex as leading causes. Anaphylaxis to NMBAs is not uncommon in patients without a known previous exposure. Quaternary ammonium ions are suggested to be the allergenic determinants in NMBAs. Commonly used household chemicals, such as toothpastes, detergents, shampoos, and cough medicines, share these same determinants with NMBAs, thereby being a contributing factor in these cases.

Clinical presentation

The etiologic diagnosis of an immediate reaction occurring during anaesthesia relies on a triad including **clinical**, **biological**, and **allergologic** evidence.



1. **Clinical:**

The initial diagnosis of anaphylaxis is presumptive as it may progress within minutes to become life-threatening. Therefore the first line of evidence for the diagnosis includes the features and severity of clinical signs as well as the timing between the introduction of the allergen and onset of symptoms. The required dosage of resuscitation drugs used also indicates the severity of the reaction.

The clinical features during anaesthesia may include:

Cardiovascular:	Tachycardia, bradycardia, cardiac arrhythmias, hypotension, cardiovascular collapse, cardiac arrest
Respiratory:	Bronchospasm
Cutaneous-mucous:	Erythema, urticaria, angioedema

(See Ring and Messmer severity scale: Grades I/II are usually not life-threatening, whereas grades III/IV are emergency situations.)

Perioperative anaphylaxis usually occurs within minutes after induction. It is primarily linked to intravenous agents. The most common presentation during severe reactions is pulselessness, desaturation and bronchospasm with difficult ventilation. Respiratory signs are exaggerated in patients that are known with underlying respiratory disease (COPD/asthma).

There are three predictive criteria for the severity of the on-going anaphylactic reaction:

- Rapid onset of reaction post-exposure to the allergen
- Cutaneous signs may be absent in rapidly progressive anaphylaxis (vasospasm of subcutaneous vascular bed during circulatory homeostasis)
- Bezold-Jarisch reflex (cardio inhibitory reflex with paradoxical bradycardia during extreme hypovolemia)

2. **Biological:**

Biochemical tests can be done *in vivo* or *in vitro.*

Primary investigation: *In vivo*

- *Histamine*: Preformed in granules of mast cells and basophils

  Early increase in plasma concentrations, plasma half-life 15-20minutes (prolonged presence in severe reactions)

- *Tryptase*: Preformed neutral serine protease in mast cells
    - Alpha-tryptase: Secreted constitutively, increased in mastocytosis (disorder of too many mast cells in the body)
    - Pro-beta-tryptase: Secreted constitutively, represent mast cell mass
    - Mature beta-tryptase: Stored in mast cell granules, reflects mast cell activation with mediator release

  Reach peak plasma concentrations 15-60 minutes, half-life approximately 2 hours. Better to compare concentrations with baseline levels taken more than 24 hours after the reaction.

Secondary investigation*: In vitro* (not available everywhere)

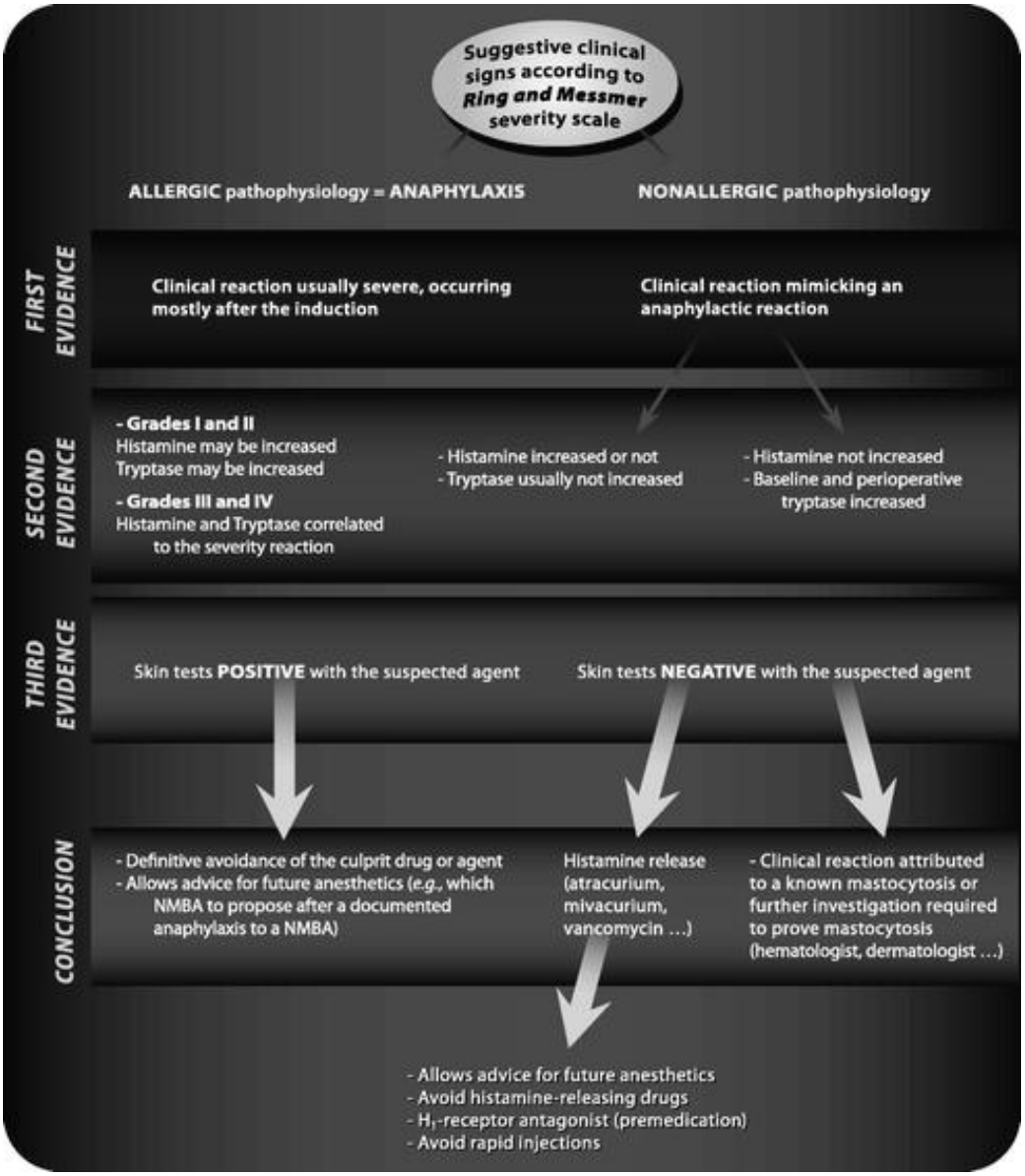
- *In vitro specific IgE assays*: Suxamethonium, Thiopental, Propofol, Amoxicillin, Cefaclor, Penicillin G/V, Latex (less sensitive than skin tests)
- *Leucocyte histamine release test*: Reliable to investigate cross-reactivity among NMBAs
- *Flow cytometric analysis*: Quantification of in vitro-activated basophils after challenge

3. Allergological/Skin Tests:

Skin tests remain the **gold standard** for IgE-mediated reactions by exposing the skin mast cells to suspected allergens. It serves to identify the culprit agent as well as the pathophysiological mechanism (allergic vs. non-allergic) to be able to suggest a safe drug alternative for future anaesthetic exposures. A 4-6 weeks delay after an anaphylactic reaction is required to avoid false-negative results.

Skin tests: Diluted or undiluted commercial solutions, Negative (saline) and positive (histamine/codeine) control

- Prick tests (PTs)
- Intradermal tests (IDTs): More sensitive/less specific, more likely to trigger systemic allergic reaction



Specific drugs

1. **Neuromuscular Blocking Agents**:

- 50-70% of perioperative anaphylaxis (suxamethonium, rocuronium)
- Previous exposure not necessary: quaternary ammonium structure in household chemicals
- Cross-reactivity between NMBAs common (60-70%)
- Benzylisoquinolones (mivacurium, atracurium) cause direct mast cell degranulation with skin changes

2. **Antibiotics**:

- Penicillins and 1st generation cephalosporins (70%) due to beta-lactam ring
- Cross-reactivity between penicillins and 1st generation cephalosporins 8-10% due to similar structure (can give 2nd/ 3rd generation cephalosporins cautiously)
- Vancomycin, if administered too quickly cause “Red Man Syndrome” due to generalised histamine release

3. **Latex**:

- Cause of 20% of perioperative anaphylactic reactions
- High risk group: Atopic patients, Food/Fruit allergies (banana, mango, kiwi...), repeated exposure to latex (repeated surgeries, spina bifida, health care workers), severe contact dermatitis
- Latex in anaesthesia: gloves, intravenous cannulas/vials, urine catheters, endotracheal tubes

4. **Other**:

- *Hypnotics*: Thiopental/Propofol rare, Etomidate/Ketamine/Benzodiazipines extremely rare
- *Opioids*: Very rare (Morphine/Codeine/Meperidine induces histamine release), cross-reactivity uncommon
- *Local Anaesthetics*: Extremely rare (Metabolite of esters, para-amino-benzoic acid may provoke IgE mediated reactions), Cross-reactivity in ester group (no cross-reactivity between ester and amide groups)
- *Colloids*: Rare, but more frequently gelatins, dextrans, albumin (Hydroxyl-ethyl starches extremely rare)
- *Dyes*: Isosulfan and patent blue rare, Methylene blue extremely rare
- *Others:* Protamine, antiseptics (chlorhexidine, povidone iodine), iodinated contrast agents (free iodine fractions)

Management of perioperative anaphylactic reactions

(Prevention in patients with previous un-investigated severe immediate reactions during anaesthesia: Regional anaesthesia with latex-free environment)

Immediate actions:

Withdraw suspected culprit

1. Discontinue anaesthetic drugs if occurred during induction
2. Maintain airway with 100% oxygen
3. Early administration of Epinephrine
4. Call for help
5. Position in Trendelenburgh position
6. Abbreviate surgical procedure/Postpone surgery

Restoration of cardiovascular homeostasis:

- Intravenous epinephrine: Titrated boluses 10-20ug (Grade I/II) or 100-200ug (Grade III/IV) +- continuous infusion (1-4 ug/min). High doses (1-3mg) with cardiac arrest during cardiopulmonary resuscitation.
- Intramuscular epinephrine 1:1000 should be administered as follows:
    - >12 years 500 ug IM (0.5 mL)
    - 6-12 years 250 ug IM (0.25mL)
    - >6 months-6 years 120 ug IM (0.12 mL)
    - <6 months 50 ug IM (0.05 mL)
- Intravascular fluids: Start rapid crystalloid/colloid boluses early to compensate for large fluid shifts due to increased vascular permeability (within 10 minutes). Adult patient may require 2-4 L of crystalloid.

Bronchospasm:

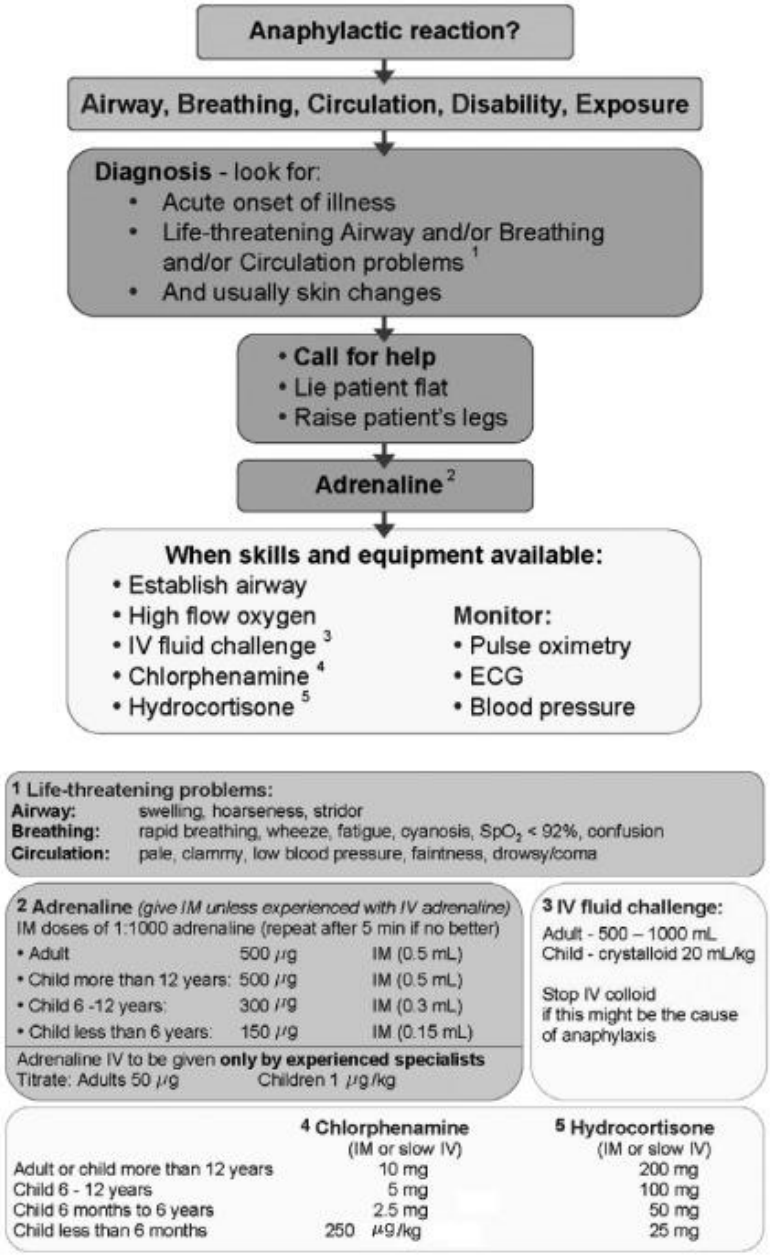
- Inhaled/Intravenous Beta2-agonists (Salbutamol/Albuterol)
- Epinephrine (Beta2-effects)
- Intravenous corticosteroids early (beneficial effects 4-6 hours later)

Additional:

- Intravenous corticosteroids and H1-receptor antagonists effects never evaluated in placebo-controlled trials

- Anaphylactic shock refractory to cathecholamines:
    - Patients on beta-blocker therapy: norepinephrine, metaraminol, glucagon
    - Vasopressin
    - Methylene blue (catecholamine- and vasopressin-resistant anaphylaxis, interfere with nitric oxide-mediated vasodilation)

(Please see RCSA algorithm at: www.resuscitationcouncil.co.za)



Summary

Anaphylaxis is a life threatening condition, more so perioperatively due to the lack of cutaneous symptoms because the patient is unconscious and draped and signs are not noticed. Adverse drug reactions and side effects are usually expected and therefore managed accordingly, whereas anaphylactic reactions are unexpected and dose independent and can occur at first exposure during anaesthesia.

Most drugs that are used during the perioperative period can cause anaphylaxis, but it is fortunately a very rare event. It is important to identify the offending agent to ensure a safe alternative for future anaesthetics by referral for skin testing post-operatively. Skin testing may unfortunately confirm the identity of the offending agent in only a minority of patients. Muscle relaxants, antibiotics and latex are the most common anaesthetic related allergens

Prevention plays an important role in the management of anaphylactic risk. Documentation during anaesthesia, referral to an allergist and appropriate labelling of the patient are essential to prevent future episodes. Patient must be fully informed and be instructed to give a thorough history.

References

1. Dewachter P, Mouton-Faivre C, Emala C. Anaphylaxis and Anaesthesia: Controversies and New Insights. Anesthesiology 2009; 111:1141-50
2. Ryder S, Waldmann C. Anaphylaxis. CEACCP 2004; Volume 4 Number 4:111-113
3. Dippenaar J, Naidoo S. Allergic reactions and anaphylaxis during anaesthesia. J Allergy Clin Immunol 2015; 28:18-22
4. Resuscitation Council of South Africa: Emergency management of anaphylaxis
5. Sampson HA, Munoz-Furlong A, Campbell RL, Adkinson NR JR et al. Second symposium on the definition and management of anaphylaxis: Summary Report-Second National Institute of Allergy and Infectious Disease/Food Allergy and Anaphylaxis Network symposium. J Allergy Clin Immunol 2006; 117:391-7

Low Flow Anaesthesia

Dr Rob Nieuwveld

UCT Dept. of Anaesthesia & Perioperative Medicine
Red Cross War Memorial Children's Hospital

INTRODUCTION

ENIGMA Many modern anaesthetists use the circle system (with CO_2 absorption) with high fresh gas flows, thus negating all the potential advantages that this system offers.
These include improved humidification and heat conservation
significant decreases in theatre and atmospheric pollution
economy in gas and volatile agent utilisation

HISTORY

1727	**S Hales** describes a breathing device with CO_2 absorption (28 years before CO_2 was identified).
1850 and 1906	**J Snow** and **F Kühn** describe and use rudimentary total rebreathing systems.
1908	Rotameters developed (Germany).
1915	**D Jackson** (Cincinnatti, USA) uses closed circuit anaesthesia on animals, after working with CO_2 absorption in submarines.
1920 to 1926	**R Waters** uses and describes his To-and-Fro closed system with canister, and determines the advantages, disadvantages and precautions needed with closed system anaesthesia.
1926 to 1930	DRÄGER patents a closed circle system. **B Sword** develops and describes the modern circle system.
1933	Cyclopropane is introduced into anaesthesia practice. Its explosive nature and expense necessitate a closed system.

WHAT ARE LOW FLOWS?

Traditionally	HIGH- Flows	> Minute Volume of fresh gas flow e.g. Mapleson DEF (T-pieces)
	MODERATE-	1,5 L min^{-1} - Minute Volume e.g. Mapleson A
	LOW-	0,5 - 1,5 L min^{-1}
	BASAL-	Supply only what is taken up by patient

These figures do not take into account the variation in patient size, weight, age, etc., and a better classification would be to compare it to predicted basal O_2 ($\dot{V}O_2$) requirements (using a factor of 2,5 for each level).

Thus	BASAL- Flow	= 1 x $\dot{V}O_2$	± 218 mL min^{-1}	(assuming 70 kg and 36,6° C)
	MINIMAL-	= 1 - 2,5 x $\dot{V}O_2$	± 218 - 550 mL min^{-1}	
	LOW-	= 2,5 - 6 x $\dot{V}O_2$	± 0,55 - 1,3 L min^{-1}	
	MODERATE-	= 6 - 15 x $\dot{V}O_2$	± 1,3 - 3,25 L min^{-1}	
	HIGH	= > 15 x $\dot{V}O_2$	> 3,25 L min^{-1}	

Predicted basal uptake of O_2 is calculated from Brody's formula.

BASIC PHYSIOLOGICAL CONSIDERATIONS

A) O_2 Requirements at basal metabolic rate (BMR)

The minimum gas flow required in a circle system is determined by the patient's basal O_2 uptake and any leaks (zero if the circuit is 100 % "tight" and monitor sampling gasses are returned to the circuit). O_2 requirements at BMR (temperature 37,6° C) may be predicted by **Brody's formula** (1945).

$$\dot{V}_{O_2} = 10{,}15 \bullet BW^{0,73}$$

Simplified to

$$\dot{V}_{O_2} = 10 \bullet BW^{0,75}$$

or

$$\dot{V}_{O_2} = 10 \bullet kg^{¾}$$

where $\dot{V}_{O_2}$ = Basal O_2 Uptake at 37,6° C in $mL\ min^{-1}$
and BW = Body Weight in kg

For each 1° *C* below 37,6° C, reduce by 10 %
e.g. 70 kg at 36,6° C = 242 - 10 % = 218
at 35,6° C = 242 - 10 % = 218 - 10 % = 196 $mL\ min^{-1}$
at 34,6° C = 242 - 10 % = 218 - 10 % = 196 - 10 % = 176 $mL\ min^{-1}$, etc.
The converse holds true for raised temperatures.
For an **adult** the basal O_2 consumption is $\pm 3\ ml\ kg^{-1}\ min^{-1}$ ($\pm 210\ mL\ min^{-1}$). At these fresh gas flow rates, it becomes imperative to continuously monitor the F_IO_2 **in** the circuit and maintain it above 0,3.

Cardiac Output at BMR has a similar relationship to Body Weight, the formula being

$$\dot{Q} = 0{,}2 \bullet kg^{0,75}$$

where $\dot{Q}$ = Cardiac Output in $L\ min^{-1}$ at 37,6° C

B) Anaesthetic agent uptake

Vaporisers may be mounted **in circuit** (VIC) or **out of circuit** (VOC), the former is dangerous unless an agent monitor is used continuously to monitor F_IAA, (the concentration in the circuit will always be higher than the dial setting). A VIC vaporiser must be of *low* resistance and *inefficient*, e.g. Goldman.

The dial setting of a VOC vaporiser will not reflect the actual delivered concentration (it will be lower) and will change with time as the patient approaches steady state conditions (may take several hours).

The amount of anaesthetic agent vapour required to produce a given MAC is a negative exponential function and is expressed as follows by **Lowe's formula**.

$$\dot{V}_{AA} = f \bullet MAC \bullet \lambda_{B/G} \bullet \dot{Q} \bullet T^{-0,5}$$

where $\dot{V}_{AA}$ = Anaesthetic vapour uptake in $mL\ min^{-1}$
$f \bullet MAC$ = Fraction of MAC desired/delivered
$\lambda_{B/G}$ = Blood / Gas Partition Coefficient of the anaesthetic agent
$\dot{Q}$ or $0{,}2 \bullet kg^{0,75}$ = Cardiac Output in $L\ min^{-1}$
and $T^{-0,5}$ or $\sqrt{T}$ = Square Root of Time in min

The square root of time concept (Severinghaus, 1954) implies that the same amount of agent is taken up in each period, each period increasing by 2 min.

Duration of anaesth (min)	0	1	4	9	16	25	36	49	64	81	100
Period no.		1	2	3	4	5	6	7	8	9	10
Duration of period (min)		1	3	5	7	9	11	13	15	17	19

N_2O uptake follows a similar pattern and the same formula applies, but a shunt factor of 0,7 is required to compensate for the rapid saturation of richly perfused organs, displacement of N_2, etc..

$$\dot{V}_{N_2O} = 0,7 \bullet f_{N_2O} \bullet 0,46 \bullet \dot{Q} \bullet T^{-0,5}$$

or, for a normal adult

$$\approx 1\,000 \bullet T^{-0,5}$$

where $\dot{V}\ N_2O$ = N_2O uptake in $mL\ min^{-1}$
and 0,46 = $\lambda_{B/G}$ of N_2O

C) Priming dose

The first dose of volatile agent must prime the breathing system and the lungs (FRC) and may be calculated from the following

$$V_{AA} = f \bullet MAC \bullet \lambda_{B/G} \bullet \dot{Q} + V$$

Where V_{AA} = Prime Dose
and V = Volume of breathing system and lungs (FRC)

This approximates the volume of anaesthetic agent taken up in the first minute (Period 1 above) and for convenience, the first dose of anaesthetic agent is usually doubled.

D) Calculation of the unit dose of liquid anaesthetic agent

Rate of uptake of anaesthetic agents (including N_2O) is an exponential process determined by $T^{-0,5}$. Initially, very high concentrations of volatile agents are required (beyond the capability of most vaporisers), and a more convenient method is direct injection of **liquid** agent into the expiratory limb. Unit Dose = $\dot{V}_{AA}$ in the 1st minute (i.e. $T^{-0,5} = 1$) and converted to liquid mL

Prediction of anaesthetic vapour uptake is calculated in *ml* of vapour, and this must be converted to *ml* of liquid anaesthetic agent using Avogadro's Hypothesis, as follows

$$mL_{vap.}\ mL_{liq.}^{-1} = (SG\ /\ MW) \bullet 22,4 \bullet (310\ /\ 273)$$

where SG = Specific Gravity
MW = Molecular Weight
22,4 = Volume of 1 Mol of gas - Avogadro's Number
and 310 / 273 = Temperature compensation in ° K for body temperature (37° C)

Results for the various agents are
Halothane = 226, Enflurane = 196, Isoflurane = 195, Sevoflurane = 182 and Desflurane = 207
Clinically 200 mL of vapour per mL of liquid is accurate enough

E) Time constant of the breathing system

All breathing systems have a time constant which determines how soon a change in fresh gas composition is reflected in the inspiratory gas. This is short with high flows and becomes increasingly longer the lower the fresh gas flow. The relationship is

$$\tau = V_S\ /\ (\dot{V}_f - \dot{V}_u)$$

where τ = Time constant in min
V_S = Total volume of the system
$\dot{V}_f$ = Fresh gas flows in $mL\ min^{-1}$
and $\dot{V}_u$ = Total gas uptake (i.e. O_2, N_2O, Anaesthetic agent and any leaks)

If a totally closed system is kept at a constant volume, then

$$\dot{V}_f = \dot{V}_u$$

and $\dot{V}_f - \dot{V}_u = 0$

$$\tau = V_S \div 0 = \infty$$

Theoretically, a completely closed system at *steady state* and basal flows will not reflect any change in fresh gas composition at equilibrium, and the fresh gas flow must be increased to make any changes.

ADVANTAGES

i) ***Economy*** in the use of gasses and anaesthetic agents is the main driving force behind the current world-wide resurgence in the use of low flow anaesthesia.
As health budgets come under strain, the economic advantages of low flow anaesthesia become obvious.
Assuming a flow reduction from 6 $L\ min^{-1}$ to 1 $L\ min^{-1}$ fresh gas flow:
- O_2 (although relatively cheap) may be reduced by a factor of 4 (2 to 0,5 $L\ min^{-1}$)
- N_2O (an expensive gas) may be reduced by a factor of 8 (4 to 0,5 $L\ min^{-1}$)
- Volatile agents vary in cost, newer agents being costly, and huge savings may be made.

However, soda-lime is relatively expensive and needs to be taken into account.

The above estimates ignore the short periods of high flow required for initial wash-in and denitrogenation. Further savings are obtained by reducing the flow even further.

ii) Increased humidity and warming of the inspired gasses
A major advantage of low flow anaesthesia, preserving mucociliary function and maintaining normothermia.
This results in less postoperative pulmonary dysfunction.
The reaction of CO_2 absorbents with CO_2 is an exothermic reaction (i.e. producing heat) and has water vapour as a by-product.
An heat moisture exchanger (HME) placed in a low flow system will not be efficient.

iii) Less pollution and easy scavenging
Less volatile agents and N_2O, implies less theatre and global pollution.
N_2O is a greenhouse gas and also depletes the ozone layer through NO.
Volatile agents (halogenated hydrocarbons) affect the ozone layer and release Cl and F.

iv) Facilitation of measuring O_2 uptake
If the breathing system is “tight” and completely closed, it may be used as a crude physiological tool for measuring gas and agent dynamics.

DISADVANTAGES

i) It is a complex system, with many components and often bulky, heavy and initially expensive.
The many couplings are all potential leaks, which is important if low flows are used.
Monitoring gases must be returned to the system.

ii) Moisture may cause unidirectional valves to stick (esp. the expiratory valve, or valves positioned near the patient), resulting in rebreathing.
There is also increased resistance to breathing, which may be a factor in spontaneous ventilation.

iii) Exhaustion of soda-lime is not always easy to detect.

iv) Cross infection may occur if the system is non-disposable as they are difficult to sterilise.
Soda-lime is, however, a very hostile environment for bacteria and viruses and no conclusive proof exists for increased nosocomial infections or ventilator associated pneumonia.
Bacterial/viral filters may be used but must be changed between cases, otherwise cross infection risk increases.
Disposable systems overcome these problems, but have a cost premium.

v) Efficient CO_2 absorption results in hypocarbia, which can only be reversed by relative hypoventilation, and this may cause increased atelectasis. It is wiser to slow the rate of ventilation than to decrease the tidal volume in this situation (which should be 6 - 8 $ml\ kg^{-1}$).

vi) Very low flow anaesthesia may result in the accumulation of other clinically important gasses. *Any* volatile substance that is produced in the body will accumulate.

- **Methane**
  Approximately 15 to 25 % of normal people produce methane gas in their bowel from intestinal organisms. Other sources include small amounts from piped gases and the atmosphere. A methane concentration above 5,4 % is combustible in O_2, but will require > 14 hr of low flow anaesthesia to accumulate
- **Acetone**
  As most surgical patients are starved (probably beyond their glycogen reserves) fat metabolism is present and this produces ketones. The volatile ketone, acetone, is exhaled and will accumulate in the system. Diabetic and cirrhotic patients have a higher rate of production. Acetone may lead to nausea, vomiting and slow emergence.
- **Inert gases**
  The only inert gas of any consequence is **nitrogen** (N_2). N_2 is neither produced nor utilised by the body, but as we live in a N_2 rich atmosphere, every tissue in the body has a partial pressure of N_2 of 78 kPa (at sea-level).
  To obtain an O_2 concentration > 21 %, and especially if we want to use N_2O, we need to denitrogenate the body tissues. This is achieved with an initial *high* fresh gas flow for 10 - 15 min, which will flush out most of the N_2 from the system, functional residual capacity (FRC) and the vessel rich group of tissues.
  However, a continuous washout of N_2 from the other tissues occurs until equilibrium is reached and this only occurs after many hours (? days). The result is a gradual accumulation of N_2 (to a maximum of 78 %) in the system and this may be *estimated* from the gas analyser:

$$N_2 = (101,2 - 6,2) - (P_IO_2 + P_IN_2O + P_ICO_2 + P_IAA)$$

or as a *rough* estimate

$$95 - (P_IO_2 + P_IN_2O)$$

as $P_ICO_2 = 0$ if soda-lime is not exhausted and P_IAA is usually small. (N.B. at sea level!)

The other inert gases e.g. Argon, Xenon, Helium, etc. cannot accumulate above the normal levels found in the atmosphere (unless the fresh gas supply has higher levels, e.g. from an O_2 concentrator).

To minimise the possible deleterious effects of these accumulated products, it is recommended to flush the system periodically at a higher flow e.g. every 1 - 2 hr. Remember to turn down the vaporiser setting!

vii) The reaction of soda-lime with some inhalational agents may produce toxic products.

- **Trichloroethylene**
  Seldom used nowadays, but produces toxic compounds when exposed to hot soda-lime ($\pm 60°\ C$) - dichloroacytelene (C_2Cl_2), a potent neurotoxin; and phosgene (COH_2), which may be lethal. Trichlorethylene is thus contra-indicated in circle systems.
- **Halothane**
  CO_2 absorbents degrade halothane to CF_2=CBrCl. This is nephrotoxic in rats, yet nephrotoxicity from halothane at any fresh gas flow is exceedingly rare in humans.
- **Sevoflurane**
  Compound A (CH_2F-O-C[=CF_2][CF_3]) is produced by the reaction of sevoflurane with CO_2 absorbents and an increased production is associated with high minute ventilation, low fresh gas flows, high absorbent temperature, Baralyme® > soda-lime and high anaesthetic concentration. Compound A has been shown to be nephrotoxic in rats and biochemical markers of renal injury were elevated in human volunteers after 8 hr at 2 $L\ min^{-1}$ fresh gas flow and 1,25 MAC with hyperventilation. However, renal function remained normal.

Sevoflurane also releases free F ions as a result of metabolism (± 3 %) and these may also be nephrotoxic (cf. enflurane metabolised ± 2 %).

Despite this potential for nephrotoxicity (Compound A and release of free F ions) no cases of sevoflurane attributable renal failure have been identified in over 2 million anaesthetics, although most of these were at fresh gas flows > 2 $L\ min^{-1}$, and the risk would increase with low flows. Many countries instituted a minimum fresh gas flow of 2 $L\ min^{-1}$ for sevoflurane.

It could be considered wise to withhold sevoflurane in patients with pre-existing renal disease and to limit the duration of exposure if low flows are used.

- **Desflurane**
  A series of patients with acute carbon monoxide poisoning, usually occurring on a Monday morning, alerted investigators to the considerable production of CO when desflurane came into contact with desiccated CO_2 absorbents. This occurs to a small degree in all CO_2 absorbents with desflurane > enflurane > isoflurane >> halothane and sevoflurane, and Baralyme® > soda-lime.
  It takes 24 - 48 hr to desiccate CO_2 absorbents with a flow of dry O_2, hence the preponderance of cases on a Monday, presumably after the O_2 flow had been left on over the weekend and routed through the CO_2 absorber.

  CarboxyHb levels > 30 % have been recorded and may be a cause for nausea, vomiting, headache and agitation. (Levels > 60 % are fatal.)
  A clue to the presence of elevated CarboxyHb is an unexpectedly high anaesthetic agent concentration reading on the agent monitor or its giving a wrong agent warning (caused by trifluoromethane produced with the CO).
  CO will accumulate if low flows are used and aggravate the problem.
  Fatal levels of CO have been produced experimentally and are possible.
  The incidence of CO exposure with anaesthesia was found to be 5 in 1 085 cases (0,46 %) in an American study.

  Other sources of CO include smokers with elevated carboxyHb levels.

viii) Extra monitoring is required if very low flows are to be used, but as these monitors are rapidly becoming standard for all anaesthesia, this cannot be claimed as a disadvantage.

MONITORING

The most important monitor is the **F_IO_2 monitor** especially if N_2O is used.
It is possible to provide hypoxic gas mixtures if the amount of O_2 does not meet the basal requirements or if the partial pressure becomes too low.

Regardless of the initial concentrations of gases in the circuit, the gas concentrations in the circuit will eventually stabilise at the concentration of the *net* flow of *each* fresh gas into the breathing circuit.
Net = what goes in (fresh gas flow) minus what goes out (patient uptake)

If 100 % O_2 or 50 : 50 O_2 / Air is used, this becomes less likely if the volume remains constant.

Volume of the system needs to be monitored and kept constant with the aid of a bag or preferably a graduated rising bellows (falling bellows generate negative pressures and may thus entrain air, but "new" generation falling bellows which prevent this have been developed).

Monitoring of inspired and expired concentrations of the various gases is an aid in maintaining constant anaesthesia levels and may indicate the changing needs of the patient, e.g. cardiac output. Agent monitors are not essential, but are becoming a "standard of care" and are a safeguard against patient awareness.

Standard monitoring, as for any other anaesthetic, is otherwise required.

HOW TO DO IT!

Low flow Anaesthesia
A standard technique for low flow anaesthesia was popularised by **Foldes** (1954).
After initial stabilisation at high flows (4 - 6 L min^{-1} for 15 - 20 min), the gas flows are reduced to 500 mL min^{-1} each of O_2 and N_2O with the vaporiser set at 1,5 MAC.

Minimal flow Anaesthesia
Standard technique described by **Virtue** (1974).
Most modern anaesthesia machines e.g. GE Datex-Ohmeda, Dräger, Blease, etc. are equipped to deliver minimal flow techniques.
Machine requirements include:

- A tight breathing system with a leak of less than 100 mL min^{-1} at a pressure of 20 hPa
- Flow meters graduated in 50 - 100 mL steps down to 100 - 200 mL min^{-1}
- Flow compensated vaporiser
- O_2 monitor in the breathing system (preferably at the Y-piece)

Initial gas flows are O_2 at 1,5 l min^{-1} and N_2O at 3,5 L min^{-1} with the vaporiser set at 1,5 MAC
N_2 is washed out and gas flows are reduced after 10 - 15 min to

O_2 = 300 mL min^{-1} and
N_2O = 200 mL min^{-1}

with the vaporiser increased to 2 MAC, e.g. isoflurane = 2,5 enflurane = 3,5 or halothane = 1,5

With these settings the F_IO_2 tends to increase from ± 0,32 to ± 0,4 whilst the F_IN_2O drops from ± 0,65 to ± 0,5 over the next 30 - 40 min. The difference of ± 10 % is caused by the accumulation of N_2.
Volatile concentrations tend to drop from ± 0,8 MAC to ± 0,65 MAC
After the initial 60 min a very slow drop occurs in the F_IO_2 whereas there is a slow rise in F_IN_2O, F_IAA and F_IN_2 and gradual reductions in N_2O and anaesthetic agent are required in accordance with the square root of time principle.
The initial period of high flows may be reduced for the new less soluble agents
The amount of overpressure required to attain initial levels of anaesthesia will decrease as the agents become less soluble i.e. halothane > enflurane > isoflurane > sevoflurane > desflurane. Conversely, the more soluble agents may be switched off earlier and the anaesthetic "coasted" towards the end of surgery.

Closed system or basal flow Anaesthesia
Anaesthesia machine requirements are more stringent:

- Flow meters must be graduated in 10 - 20 mL steps
- Leaks must be minimal
- Tables, calculator or computers to determine gas uptake requirements

Vaporisers may be used, but in the initial stages cannot supply the concentrations required.
In circle vaporisers (VIC) would overcome the problem but are dangerous. A more logical approach is to utilise direct injection of anaesthetic agent into the expiratory limb and this may be achieved by:

i) Manually, using a calculated unit dose and the square root of time principle
ii) Programmable exponential syringe pumps
iii) Sophisticated closed-loop electronic controllers with agent injectors

After the usual high flow washout of N_2 with 100 % O_2 and a normal IV induction, O_2 flows are reduced to Basal requirements and N_2O at maximum possible untill O_2 monitor shows a F_IO_2 of 0,4 (15 - 30 s).
N_2O flows are reduced to the rate required to maintain a constant volume in the breathing system as observed at the bellows or bag with the spill valve completely closed. Direct injection of volatile agent is commenced as described.
If a TEC vaporiser is used, it is set at the maximum (5 %) for 9 min and then reduced using the Square Root of Time (this is not possible with enflurane)

The above techniques are much simplified with the use of anaesthetic agent analysers. The end-tidal concentration will reflect MAC.

References / Recommended reading

1. Awati MN, et al. *Low Flow Anaesthesia*. Journal of Evidence Based Medicine and Healthcare 2014; 1: 1150-1162.
2. Nunn G. *Low-flow anaesthesia*. Contin Educ Anaesth Crit Care Pain 2008; 8: 1-4. doi: 10.1093/bjaceaccp/mkm052
3. Baxter AD. *Low and minimal flow inhalational anaesthesia.* Can J Anaesth 1997; 44: 643-652
4. Nel MR, Ooi R, Lee DJH, Soni N. *New agents, the circle system and short procedures.* Anaesthesia 1997; 52: 364-367
5. Eger EI II, et al. *Nephrotoxicity of Sevoflurane vs Desflurane Anesthesia in Volunteers.* Anesth Analg 1997; 84: 160-168
6. Lowe HJ, Ernst EA. *The quantitative practice of anaesthesia.* Williams and Wilkins, Baltimore, 1981.

Inhalational agents

Dr Ernest Welch

Private Practice
Honorary lecturer- Wits University

I have not provided notes on this topic as this is a large core pharmacology topic for all anaesthetists that requires detailed knowledge and understanding. I fear this cannot be provided in a summary of a few pages. I would direct you to a comprehensive chapter I have written on the topic that can be found in:
Applied Pharmacology for Anaesthesiology and Critical Care.
Edited by Analee Milner and Ernest Welch.
Published by Medpharm publishers in 2012.
Available from www.milnerandwelch.co

I think that the areas that need to covered and understood on this topic are:

1. A history of the development of the agents we use today, and how they have been improved over time.
2. A classification of the agents
3. The mechanisms of action that have been proposed for the inhalational agents.
4. The relationship between the structure and function of all the agents.
    - you must know how to draw these agents' chemical structures.
    - Isomerism, chirality and enantiomers
5. The pharmacokinetic principles of these agents which are unique as they are not administered via conventional routes.
    - These principles include
        - Uptake into the alveolus
        - The establishment of concentration gradients between lung, blood and brain.
        - Metabolism and elimination
        - The principle and application of MAC and its derivatives.
6. The pharmacodynamic principles and normal effects on organ systems of all the agents.
7. The side effects and special circumstances associated with these agents in particular those unique to anaesthesia:
    - Malignant hyperthermia
    - Halothane hepatitis
    - Renal failure
8. The technique and principles of low-flow anaesthesia
9. The principles of what an ideal agent would represent.
10. The properties of all the individual inhalational agents: (even those you may not have used in routine clinical practice)
    - Halothane
    - Enflurane
    - Isoflurane
    - Desflurane
    - Sevoflurane
    - Nitrous oxide
    - Ether
    - Methoxyflurane
    - Xenon

Notes page

Non-steroidal anti-inflammatory drugs (NSAIDs)

Dr Ernest Welch

Private Practice
Honorary lecturer- Wits University

Non-steroidal anti-inflammatory drugs (NSAIDs) are a group of unrelated chemical compounds that have analgesic, anti-inflammatory and antipyretic effects. The similarity in therapeutic actions and side effects is due to common mechanisms of action and as a result they can be studied as a single class of drugs. The understanding of the NSAIDs, their effects and controversies is dependent on knowledge of the COX (cyclo-oxygenase) enzyme system.

Mechanism of action

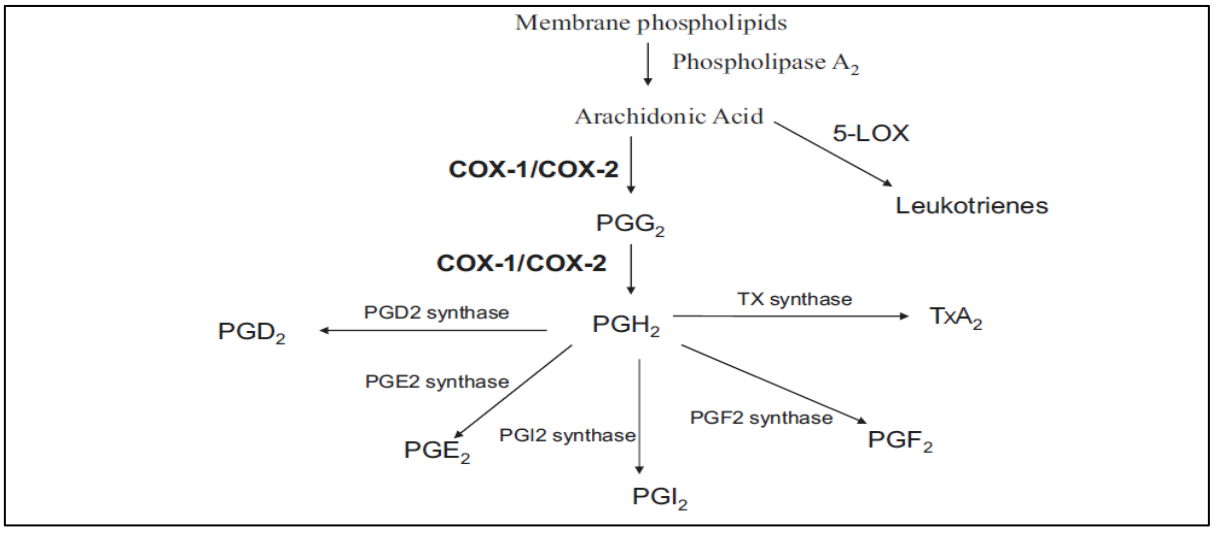
Basic physiology of the COX pathway:

1. Fatty acid metabolism results in the production of prostaglandins (PG) via the COX pathway.
2. PGs mediate: inflammation, pain, pyrexia, cell mitosis and neuromuscular function.
3. All NSAIDs inhibit cyclo-oxygenase (COX).

Formation of prostaglandin

- Arachidonic acid (AA) is a phospholipid fatty acid found in cell membranes that is released by a variety of stimuli particularly membrane damage.
- Cyclo-oxygenase (COX) and lipoxygenase (LOX) enzymes convert Arachidonic acid (AA) to lipid mediator's PG and leukotrienes (also known as the eicosanoids)
- The two COX isoforms (COX-1 and COX-2) catalyse AA to PG and thromboxane (TxA)
- Initially COX converts AA to prostaglandin G_2 (PGG_2), and then converts PGG_2 to prostaglandin H_2 (PGH_2)
- PGH_2 is converted to 5 active forms of PG
    - prostaglandin D_2 (PGD_2),
    - prostaglandin E_2 (PGE_2)
    - prostaglandin $F_{2\alpha}$ ($PGF_{2\alpha}$)
    - prostacyclin (PGI_2)
    - thromboxane A_2 (TxA_2)
- These 5 prostanoids act as secondary messengers mainly on G protein-coupled receptors.

THE SYNTHESIS OF PROSTOGLANDINS USING CYCLO-OXYGENASE



S. Bacchi, P. Palumbo, A. Sponta and M.F. Coppolino. Clinical Pharmacology of Non-Steroidal Anti-Inflammatory Drugs: A Review, Anti-Inflammatory & Anti-Allergy Agents in Medicinal Chemistry, 2012, 11, 52-64

COX has 2 main isoforms. The gene for COX-1 is located on chromosome 9 and COX-2 on chromosome 1.

COX-1 is constitutively expressed in most tissue types and the prostanoids produced by this isoform generally mediate 'housekeeping' functions: protection of the gastric mucosa, regulation of renal blood flow and platelet aggregation.
It can be induced during cell differentiation and angiogenesis
Platelets only express COX-1.

COX-2 is highly inducible (but constitutively expressed in the brain, spinal cord and kidneys.)
COX-2 is induced by: inflammatory conditions in response to cytokines (interferon, TNF, IL1), hormones, growth factors and hypoxia, and is found in tumour endothelium.

COX-3 has been described and is a variant of COX-1 that appears to have no prostaglandin producing activity.

Pharmacology of NSAIDS
NSAIDs act by blocking COX and reducing the synthesis of PG resulting in a decrease in inflammation, pain and fever.

- The anti-inflammatory action is due to the decrease in vasodilation and oedema (PGE_2, PGI_2).
- PGE_2 and PGI_2 inhibition results in central and peripheral analgesic effects.
- The antipyretic effect results from blocking production of PGE_2 modulated hypothalamic thermoregulation.

COX-2 inhibitors (COXIBs) were developed to maintain the desirable NSAID effects of analgesia, anti-inflammation and antipyretic without inhibiting the homeostatic effects of COX-1 on the GIT, kidneys and platelets.
Unfortunately, it's now known that these 2 enzyme systems are not exclusive and the COXIBS still have effects on the homeostatic systems and the side effects remain. In addition, following long term studies the COXIBs have demonstrated cardiovascular risks that have resulted in the withdrawal of rofecoxib and warnings about their use in patients at risk of cardiovascular disease.

Classification of NSAIDs
As a group of different chemical compounds, logically it's easy to classify the NSAIDs according to their drugs class; but this is of little clinical relevance and classification by means of their clinical and side effect profiles as a result of the varying amounts of COX inhibition is far more practical.

CLASSIFICATION OF NSAIDs BY CYCLO-OXYGENASE (COX) INHIBITION

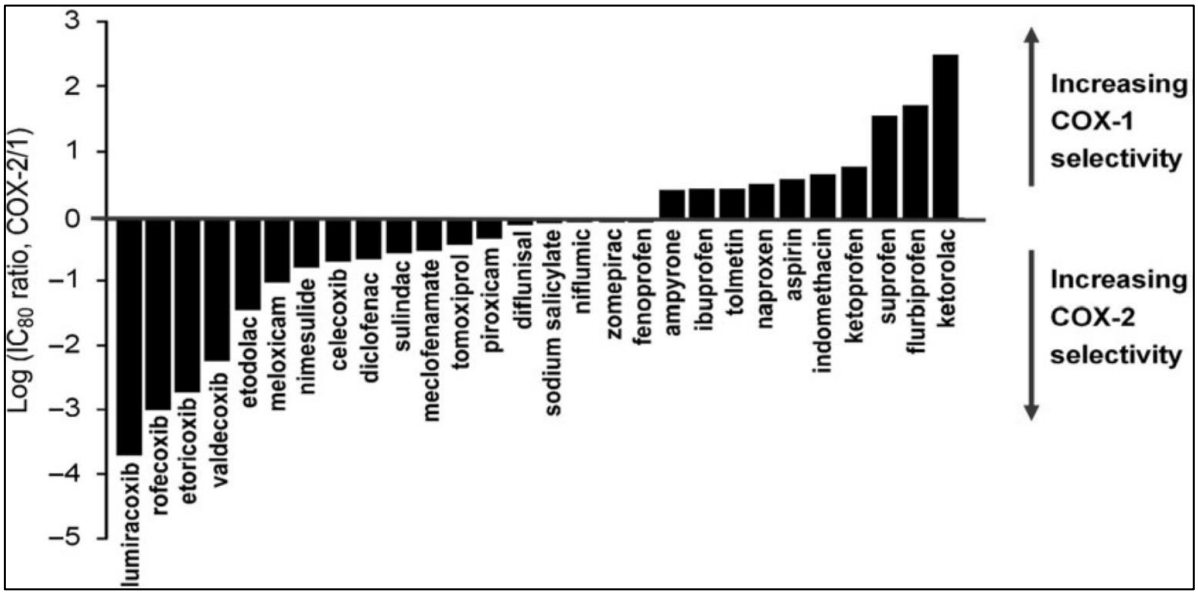
Group	COX inhibition	Example
Group 1	***Poorly selective.*** • Fully inhibit both COX-1 and COX-2 • Less than 5 times COX-2 selectivity	Ibuprofen, diclofenac, aspirin, piroxicam, naproxen
Group 2	***Preferential COX-2 selective*** • Inhibit both COX-1 and COX-2 • 5 to 50 times COX-2 selectivity	Celecoxib, meloxicam, nimesulide, etodolac
Group 3	***Predominantly COX-2 selective*** • Weak COX-1 inhibition • Greater than 50 times COX-2 selectivity	Rofecoxib
Group 4	***Weak inhibitors of both COX-1 and COX-2***	Sodium salicylate, nabumetone

CLASSIFICATION OF NSAIDs BY CHEMICAL COMPOUND / DRUG CLASS

Drug class	Example	Comments
Salicylic acid derivates	aspirin	
Indole and indene acetic acids	indomethacin, etodolac, sulindac	
Hetero-aryl acetic acids	diclofenac, ketorolac, tolmetin	
Aryl-propionic acids	ibuprofen, ketoprofen, flurbiprofen, naproxen, fenoprofen, oxaprozin, aceclofenac, fenclofenac	
Anthranilic acids	mefenamic acid, meclofenanic acid	Also called fenamates
Enolic acids	piroxicam, tenoxicam, meloxicam	Also called oxicams
Alkanones	nabumetone	Not available is SA (Relafen®)
Pyrazolidinediones	pheynylbutazone, oxyphenylbutazone	
Diarylheterocycles	lumiracoxib > rofecoxib > etoricoxib > valdecoxib > parecoxib > celecoxib	Selective COX-2 inhibitors In rank order of COX-2 inhibition
Para-aminophenol derivates	paracetamol (acetaminophen)	Paracetamol is classified as a NSAID despite having little anti-inflammatory activity

Meta-analysis has shown that there is no significant analgesic difference between the classes of NSAIDs or between the selective COX-2 inhibitors and poorly-selective agents in all forms of arthritis and orthopaedic pain.

RANKING OF CYCLO-OXYGENASE (COX) SELECTIVITY



Morten Schmidt, Morten Lamberts, Anne-Marie Schjerning Olsen et al. Cardiovascular safety of non-aspirin non-steroidal anti-inflammatory drugs: review and position paper by the working group for Cardiovascular Pharmacotherapy of the European Society of Cardiology. European Heart Journal – Cardiovascular Pharmacotherapy (2016) 2, 108–118

Pharmacokinetics of all NSAIDs are generally very similar:

- Lipid-soluble weak acids.
- Most are completely absorbed from the GIT.
- Have little first-pass hepatic metabolism.
- Highly protein bound with a small volume of distribution.
- Metabolised by CYP3A and CYP2C and /or glucuronidation
- Half-lives vary from less than 2 to 8 hours.

Adverse effects

Gastro Intestinal Tract effects

- 1-3% of patients on chronic NSAIDs develop gastrointestinal bleeding.
- PGs inhibit H^+ secretion and promote mucous production. NSAIDs block this promoting gastric erosion, ulceration and bleeding.
- COX-2 selective inhibitors theoretically have less GIT complications than non-selective NSAIDs, but this has not been conclusively shown in large trials with chronic use.
- The risk of GIT symptoms increases if patients are on concomitant aspirin. (CLASS, TARGET and VIGOR trials)

Cardiovascular

Long term trials on the COX-2 specific inhibitors showed an increased risk of 1.5 to 5 times of cardiovascular complications when compared to placebo. Non- selective NSAIDs except aspirin also have an increased risk. The safest drug in this setting appears to be naproxen.

Myocardial infarction, cerebrovascular accident and thrombo-emboli:

- COX-2 inhibition causes suppression of prostacyclin (PCI_2), which is a potent vasodilator and prevents platelet aggregation helping protect endothelial cells during shear stress and inhibiting smooth muscle cell proliferation.
- Platelets contain only COX-1, which produces thromboxane A2—a potent vasoconstrictor, platelet aggregator and pro-thrombotic.
- Selective COX-2 inhibition therefore results in unopposed COX-1 effects - promoting thrombosis.
- COX-2 inhibitors also accelerate atherosclerosis and raise blood pressure.
- COX-2 inhibition is associated with increased infarct size, myocardial wall thinning and myocardial rupture post myocardial infarction.
- Patients with cardiovascular disease or risk factors (such as hypertension, hyperlipidaemia, diabetes mellitus, or smoking) have an increased risk of thromboembolic events and NSAIDs should be avoided.

Arrhythmias

- COX-2 produced prostacyclin acts as an endogenous antiarrhythmic agent via inhibition of epicardial sympathetic nerve activity.
- NSAIDs also elicit proarrhythmic effects through fluid retention and electrolyte disturbances.

Cardiac failure

- NSAIDs can cause fluid retention and elevate blood pressure doubling the risk of heart failure.

Renal

- NSAIDs are associated with renal toxicity
- PGs are involved in controlling renin release, regulating vascular tone, and controlling tubular function.
- COX inhibition produces changes in fluid and electrolyte control
- COX-1 controls renal blood flow and perfusion.
- COX-2 is involved in diuresis and electrolyte balance.
- COX-2 inhibition is associated with increased potassium levels
- Inhibition of these enzyme systems can have profound effects on renal function.

Hepatic

- Liver toxicity from paracetamol is well known. It is the result of decreased glutathione required for metabolism of paracetamol.
- Diclofenac, nimesulide and sulindac, lumiracoxib have all been associated with liver dysfunction and failure.

Bleeding

- By inhibiting the formation of TxA_2 NSAIDS and aspirin decrease platelet function.

- Evidence of bleeding with NSAIDs is of concern only when use in combination with other anti-coagulants.
- Aspirin binds platelets irreversibly, the anticoagulant effect of other NSAIDs is usually reversible and of much shorter duration.

Bone healing

- The evidence around bone healing and NSAIDS is scanty, contradictory and largely opinion based.
- Evidence exists for delayed fracture healing but NSAIDs are also shown to make no difference.
- The potential mechanisms of delaying healing have not been described.
- The current recommendations are to avoid NSAIDs in case where bone healing may be a concern.

Newer uses:

NSAIDs and malignancy

- Evidence of the protective effect of NSAIDs in malignancy (colon, breast, prostate and pancreas) are being described.
- This is probably on the basis of PGE_2 inhibition which is associated with angiogenesis and apoptosis in cancer cells.

NSAIDs and neuromuscular disease

- COX-1 and COX-2 are found in the central nervous system of patients with ischemic and traumatic brain injury as well as in neurodegenerative diseases such as Alzheimer's and Parkinson's disease.
- COX-mediated neuroinflammation is a component in neuronal degeneration.
- Use of NSAIDs to attempt to reverse or arrest this degeneration is still experimental.

These notes are designed as a summary for the FCA 1 refresher course and have not been peer reviewed.

References

1. *S. Bacchi, P. Palumbo, et al. Clinical Pharmacology of Non-Steroidal Anti-Inflammatory Drugs: A Review, Anti-Inflammatory & Anti-Allergy Agents in Medicinal Chemistry, 2012, 11, 52-64*
2. *Morten Schmidt, Morten Lamberts, Anne-Marie Schjerning Olsen et al. Cardiovascular safety of non-aspirin non-steroidal anti-inflammatory drugs: review and position paper by the working group for Cardiovascular Pharmacotherapy of the European Society of Cardiology. European Heart Journal – Cardiovascular Pharmacotherapy (2016) 2, 108–118*
3. *Ippokratis Pountos, Theodora Georgouli, et al. Do Nonsteroidal Anti-Inflammatory Drugs Affect Bone Healing? A Critical Analysis. Scientific World Journal. 2012; 2012: 606404.*

The novel oral anti-coagulant drugs (NOACs)

Dr Felipe Montoya-Pelaez

UCT Dept of Anaesthesia & Perioperative Medicine

Introduction

Due to changes and improvements in treatment protocols for conditions where a high risk of thrombotic or thromboembolic phenomena exist, it is becoming increasingly common for anaesthetists to care for patients taking a variety of anticoagulant drugs in the perioperative period. Risks of anaesthesia and surgery in this group of patients are self-evident, including the possibility of complications related to the withdrawal or withholding of these drugs pre and post-operatively.

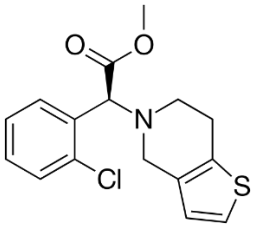
Anticoagulant pharmacopoeia was, up until the early 2000s, restricted to warfarin, heparin and its derivatives and aspirin. The limitations of these agents (inter-patient variability in response, the need for ongoing monitoring, thrombocytopaenia and the need for parenteral administration) have necessitated the development of newer agents. Presently, a new variety of drugs, effective orally and targeting novel areas in the coagulation cascade, are in widespread use for conditions such as prophylaxis of thrombo-embolic phenomena in inherited and acquired conditions, atrial fibrillation, patients with mechanical heart valves and prophylaxis and treatment of DVT and pulmonary embolism.

This review will discuss only *orally* ingested drugs introduced into clinical practice within the past 15 years:

1) **Anti-platelet agents** – *P2Y12 receptor antagonists*: Clopidogrel, Prasugel and Ticagrelor
2) **Anticoagulants** – *Direct thrombin inhibitors*: Dabigatran
   Factor Xa inhibitors: Rivaroxaban, Apixaban and Fondaparinux

1) Antiplatelet agents – P2Y12 receptor antagonists

Clopidogrel



Clopidogrel bisulfate, a second generation thienopyridine derivative, is a relatively novel anti-platelet agent indicated for the treatment and prevention of arterial thrombosis in patients at high risk of acute myocardial syndromes (ACS), ischaemic stroke and peripheral vascular disease. Clopidogrel may be used as a single agent in preventative treatment of arterio-thrombotic events or as dual anti-thrombotic therapy with aspirin in patients who have had an acute myocardial infarct, or have undergone a percutaneous coronary intervention (PCI) with or without stent insertion.

Clopidogrel is a prodrug. Only a small portion (15%) is metabolised to its active thiol metabolite in a 2 step metabolic pathway involving cytochrome P450 enzymes in the liver (specifically CYP2C19, CYP1A2, CYP2B6, CYP2C9 and CYP3A4). The rest is hydrolysed by carboxylesterase 1(CES-1) to an inactive carboxylic acid derivative that accounts for 85% of clopidogrel-related compounds. Clopidogrel irreversibly inhibits the binding of ADP to the P2Y12 receptor on the platelet membrane. Two receptors, P2Y12 and P2Y1 when activated by ADP, result in the activation of the GPII/IIIa glycoprotein complex which initiates platelet aggregation. Activation via P2Y12 is the more potent pathway **(*figure 1*).** This effect lasts for the lifetime of the platelet. Clopidogrel also has anti-inflammatory effects, achieved by decreases in the expression of platelet-dependent inflammatory markers (C-reactive protein etc.)

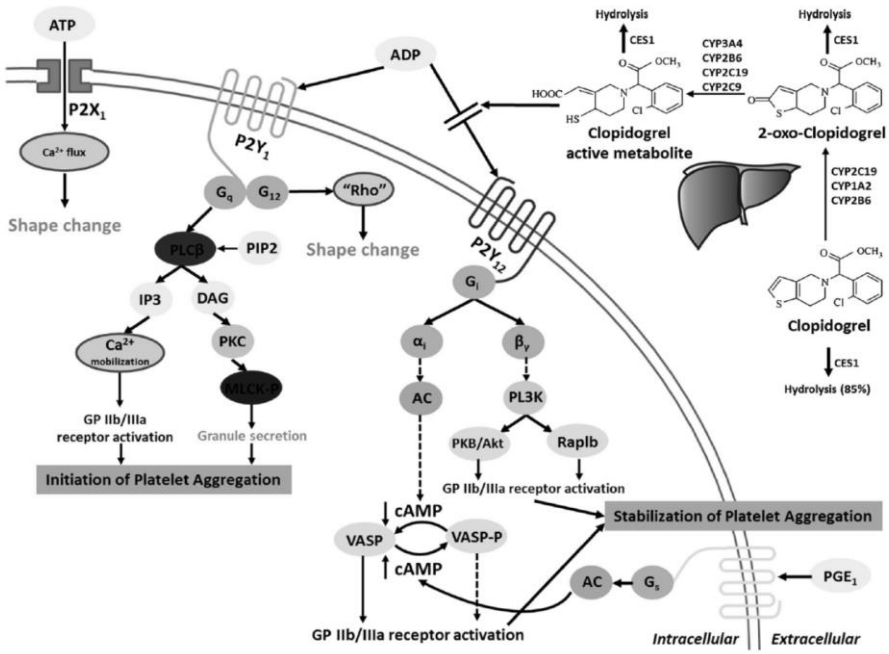


Figure 1

Dose-dependent inhibition of platelet aggregation can be seen 2 hours after single oral dose. Repeated doses of 75 mg per day inhibit ADP-induced platelet aggregation on the first day, and inhibition reaches steady state between Day 3 and Day 7. At steady state, the average inhibition level observed with a dose of 75 mg Plavix per day was between 40% and 60%. Peak inhibition may be achieved within 2 hours by administering a loading dose of 300-600mg followed by 75mg to 150mg daily (for patients with ACS undergoing PCI with or without stenting). Platelet aggregation and bleeding time gradually return to baseline values after treatment is discontinued, generally in about 5 days. Oral administration leads to rapid absorption (>50% of administered dose) which may be enhanced by food intake. After a 75mg oral dose of clopidogrel, 50% is excreted in the urine and 46% in faeces over a period of about 5 days. It has a half-life of approximately 6 hours. The half-life of the active metabolite is about 30 minutes.

Like aspirin, clopidogrel shows considerable interpatient variability in its pharmacokinetics and dynamics. This is clinical relevant because non-responders are at increased risk of thrombotic events whilst on apparently adequate dosing regimens. The prevalence of clopidogrel *resistance* varies widely, reported at between 4 and 30%. Potential mechanisms of resistance can be classified as extrinsic or intrinsic (table 1).
Pharmacogenetic factors also play a major role in determining whether a patient is a good or poor drug responder. Genetic polymorphisms that affect the pharmacology of clopidogrel include mutations of enzymes involved in its metabolism, the most important being those involving the CYP2C19 enzyme. Genetic polymorphisms affecting the P2Y12 receptor are also thought to result in degrees of resistance to clopidogrel but studies looking into this phenomenon have been inconclusive.

Table 1: Potential mechanisms of Clopidogrel resistance

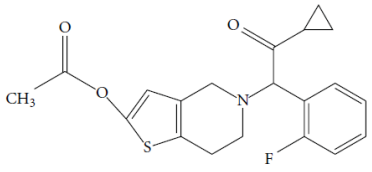
Extrinsic mechanisms:

1. Patient non-compliance
2. Inappropriate dosing
3. Drug-drug interactions involving CYP3A4 and CYP2C19

Intrinsic mechanisms:

1. Genetic variables
    a. Polymorphisms of P2Y12 receptor
    b. Polymorphisms of CYP2C19
2. Increase release of ADP
3. Alternate pathways of platelet activation
    a. Failure to inhibit catecholamine-mediated platelet activation (epinephrine)
    b. Greater extent of P2Y1 – dependent platelet aggregation
    c. Up-regulation of P2Y12 -independent pathways (thrombin, thromboxane A2, collagen)

Prasugrel

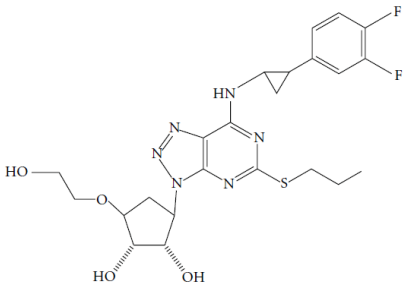


Prasugrel

Prasugrel is a third generation thienopyridine. After oral ingestion, it is hydrolysed in the gasto-intestinal tract (but not as extensively as clopidogrel) to an intermediary metabolite which is then activated in the liver by the cytochrome P450 system in a one step process. Prasugrel binds irreversibly to the P2Y12 receptor resulting in up to 90% inhibition of platelet aggregation within 2 to 4 hours of administration. The active metabolite has a half-life of approximately 7 hours. Recovery of platelet function takes longer than clopidogrel with 75% of patients returning to baseline platelet reactivity within 7 days. Variability of response to prasugrel is reduced compared to clopidogrel with CYP variants having little influence on the active metabolites of prasugrel.

Equipotent doses of prasugrel compared with clopidogrel in dual therapy with aspirin (TRITON-TIMI 38 trial) showed a rate of primary outcomes (death from cardiovascular causes, non-fatal MI and non-fatal stroke, with a safety end-point of major bleeding) that was lower at 15 months for prasugrel (9.9% vs 12.2%, p,0.001)); a reduced rate of in-stent thrombosis (1.1% vs 2.4%, p<0.001) but a higher rate of bleeding (2.4% vs 1.8%, p=0.03).

Ticagrelor



Ticagrelor

Ticagrelor is a P2Y12 inhibitor that *reversibly* binds to a site on the receptor distinct from ADP. It acts directly on the receptor site without requiring metabolic activation. It undergoes enzymatic degradation to at least one active metabolite that is as potent as its parent compound. Maximum plasma concentration and platelet inhibition occurs 1 to 3 hours after administration. The drug has a rapid offset of action with a half-life of 12 hours requiring twice daily dosing.

Like prasugrel, it exhibits greater platelet inhibition (60%) than clopidogrel. NICE guidelines based the PLATO trial recommends ticagrelor as a treatment option for patients with ACS or patients with unstable angina who are treated with PCI. The trial, comparing ticagrelor to clopidogrel in dual therapy with aspirin in 18000 patients with ACS demonstrated a lower mortality at 12 months for ticagrelor (4.5% vs 5.9%, p<0.001). There was a similar risk of major bleeding in both groups but the rate of both fatal and non-fatal intracranial haemorrhage was higher in the ticagrelor group.

Table 2 NICE recommendations for P2Y12 receptor inhibitors as dual therapy with aspirin

Drug	Indication	Dose	Duration of therapy
Prasugrel	Percutaneous Coronary Intervention – Bare Metal Stent	60mg load, then 10mg od 5mg od if >75y or < 60kg	4-6 weeks
	Percutaneous Coronary Intervention – Drug Eluting Stent	As above	12 months
Ticagrelor	Acute Coronary Syndrome requiring immediate Percutaneous Coronary Intervention	180mg load, then 90mg bd	Up to 12 months
Clopidogrel	ACS with or without PCI	300-600mg load, then 75mg od	Initially 4 weeks, Up to 12 months

2) Oral anticoagulants

For more than 60 years vitamin K antagonists (VKAs) such as warfarin have been the mainstay of oral anticoagulation for the prevention of atrial fibrillation associated stroke as well as for the treatment and prevention of deep venous thromboembolism and associated pulmonary embolism. Despite its efficacy in preventing stroke and venous hromboembolism, management of anticoagulation with warfarin can be difficult due to its slow onset and offset of action, narrow therapeutic range (requiring frequent monitoring and dose adjusting) and interactions with a variety of drugs that may enhance or hinder its anticoagulant effect. Consequently, new oral anticoagulants with specific advantages and disadvantages compared to VKAs (table 3), have been introduced that aim to address these challenges. These include the direct thrombin inhibitor dabigatran and the factor Xa inhibitors rivaroxaban, apixaban and edoxaban. *Fondaparinux* belongs to the same class of drug but is administered parenterally and will not be discussed further.

Table 3

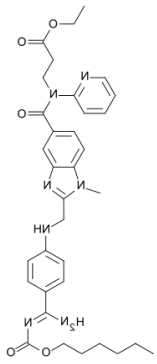
Advantages of novel anticoagulants	Disadvantages of novel anticoagulants
• Rapid onset of activity • Short half-lives • Similar or improved efficacy • Fewer drug interactions • Fixed dose guidelines • No requirement for monitoring	• No antidote • No widely available measure of activity • Limited knowledge and experience • Drug interactions • Effect of renal impairment on pharmacokinetics

In phase III clinical trials involving patients with AF, the direct acting oral anticoagulant agents (DOACs) were at least as effective in preventing stroke and systemic embolic events as warfarin with a reduced all-cause mortality and incidence of intra-cranial haemorrhage. However, non-fatal GI bleeding was increased relative to warfarin. This may be attributed to the lower absorption of DOACs across the GI mucosa compared to warfarin, which may increase the rate of bleeding of susceptible lesions in the GI tract. In the treatment of venous thromboembolism (DVT and PE) DOACs are just as effective as standard therapy with lead-in LMWH or unfractionated heparin followed by oral coumadin drugs. Rates of bleeding were lower with DOACs.

Unlike the VKAs, the new oral anticoagulants do not presently have specific reversal agents. Non-specific agents like prothrombin complex concentrates and factor VIIa are effective in a dose dependent fashion. Dialysis may be helpful for removal of dabigatran but not for the factor Xa inhibitors.

Several potential antidotes are currently in development. Andexanet alfa is a modified recombinant factor X produced from Chinese hamster ovary cells. It lacks intrinsic procoagulant effects and is able to bind direct factor Xa inhibitors without interfering with the coagulation cascade.
Another promising antidote under development is aripazine. It binds non-covalently to and inhibits the activity of both direct and indirect anticoagulants, including both oral and parenteral agents.

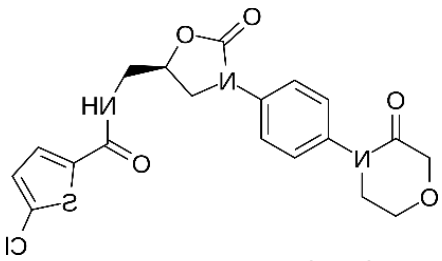
Dabigatran



Dabigatran etixalate is a competitive, highly specific direct thrombin inhibitor. It is administered as an inactive drug and converted to its active form by ester hydrolysis.
It has a rapid onset of action (1-2 hours) and has a half-life of 12-17 hours. It is 35% bound to plasma proteins with 80% of an administered dose excreted by the kidneys. Recommended dosages are 150mg BD and 110mg BD for patients over 80 years old or those at a high risk of bleeding. Lower doses are recommended for patients with reduced renal function.
Dabigatran etexilate is a substrate for the transmembrane transporter P-glycoprotein (P-gp). Drugs that inhibit P-gp (e.g verapamil and amiodarone) increase the bioavailability of dabigatran

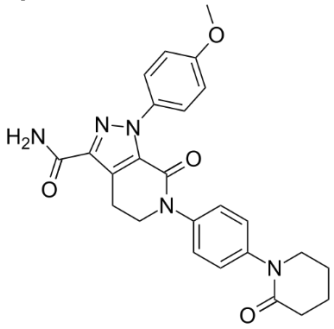
Rivaroxaban

Rivaroxaban is a direct, competitive and dose dependent inhibitor of factor Xa. It is rapidly absorbed and reaches peak plasma concentrations within 2-4 hours and has a half-life of 9-13 hours. It is highly protein bound (up to 95%). 35% of an administered drug is cleared by the kidney, the rest is metabolised in the liver by the cytochrome P450 enzyme complex (CYP3A4 and CYP2J2). Treatment with P450 iso-enzymes and P-gp inhibitors such as itraconazole and voriconazole is contraindicated due to an increased risk of bleeding. Enzyme inducers such as carbamazepine, phenytoin and St John's Wort are liable decrease anticoagulant effect. Recommended dosing for stroke prevention in AF is 20 mg daily with VTE prophylaxis at 10mg daily. Renal dysfunction mandates dose reduction. Food enhances absorption.



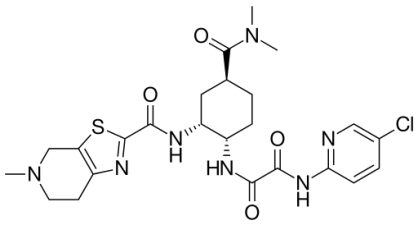
Rivarobaxan

Apibaxan



Apibaxan is a direct, competitive and selective inhibitor of factor Xa and, like the dabigatran and rivaroxaban, it is approved for stroke prevention in AF and VTE prophylaxis in orthopaedic surgery. It is well absorbed, achieving peak plasma concentrations in 1-4 hours. Plasma half-life is approximately 12 hours. 50% is faecally excreted, 25 percent via the kidney. Apixaban is a substrate for both P-gp and CYP3A4. It is administered orally in a dose of 5mg twice daily with half this regimen recommended for patients aged over 80 years, patients weighing less than 60 kg and those with a serum creatinine of > 133µmol/L.

Edoxaban



Edoxaban

Edoxaban is another direct factor inhibitor. Peak plasma concentrations of the drug occur 1-2 hours following oral administration. It has a bioavailability of 62% and is 55% protein bound. Most of the drug remains unchanged in the plasma although there is minimal metabolism via hydrolysis, conjugation and oxidation by CYP3A4. 50% of the drug is eliminated unchanged in faeces with approximately 50% the urine. Over70% of the drug is eliminated unchanged with an elimination half-life of 9-11 hours. Dosing ranges from 60mg daily to 30mg daily in patients with reduced creatinine clearance (less than 50ml/min). Interestingly edoxaban is not recommended in patients with creatinine clearances of >95mls/min due to an increased incidence of stroke compared to patients on warfarin.

Although edoxaban appears to have fewer drug interactions than other direct Xa inhibitors it still has potentially clinically important interactions with inhibitors or inducers of the P-gp efflux transporter.

Table 4: Pharmacology of the novel oral anti-coagulants

Drug	Dabigatran	Rivaroxaban	Apixaban	Edoxaban[a]
Mechanism	Direct thrombin inhibitor	Direct factor Xa inhibitor	Direct factor Xa inhibitor	Direct factor Xa inhibitor
Pro-drug	Yes	No	No	No
Bioavailability, %	6%	66% without food up to 100% with food	50%	62%
Half-life, h	12–17 h	5–9 h (young) 11–13 h (elderly)	12 h	9–11 h
Time to maximum plasma concentration	0.5–2	2–4 h	1–4 h	1–2 h
Renal excretion	80%	35%	25%	50%
Liver metabolism	No	Yes	Yes	Minimal
Gastrointestinal tolerability	Dyspepsia	No problem	No problem	No problem
Absorption with food	No effect	+39% more	No effect	6–22% more
Intake with food?	No	Mandatory	No	No official recommendation
Dosing	Twice daily	Once daily	Twice daily	Once daily

Management of patients presenting for elective surgery

There is a paucity of evidence available to guide the management of these new medications around the time of surgery. Because of the large inter-patient variability in drug handling and the unpredictability of residual anticoagulant effect, decisions in the perioperative period must be tailored to individual patients balancing the risk of significant bleeding against that of thromboembolic events. If the patient is undergoing a procedure with a low or minor risk of bleeding then it is reasonable to continue treatment; conversely, if a patient is at a low-risk of thrombosis undergoing a procedure with a high-risk of bleeding then the treatment can be suspended perioperatively. Obviously, the problem arises with patients who are at high-risk for thrombosis undergoing high bleeding risk procedures.

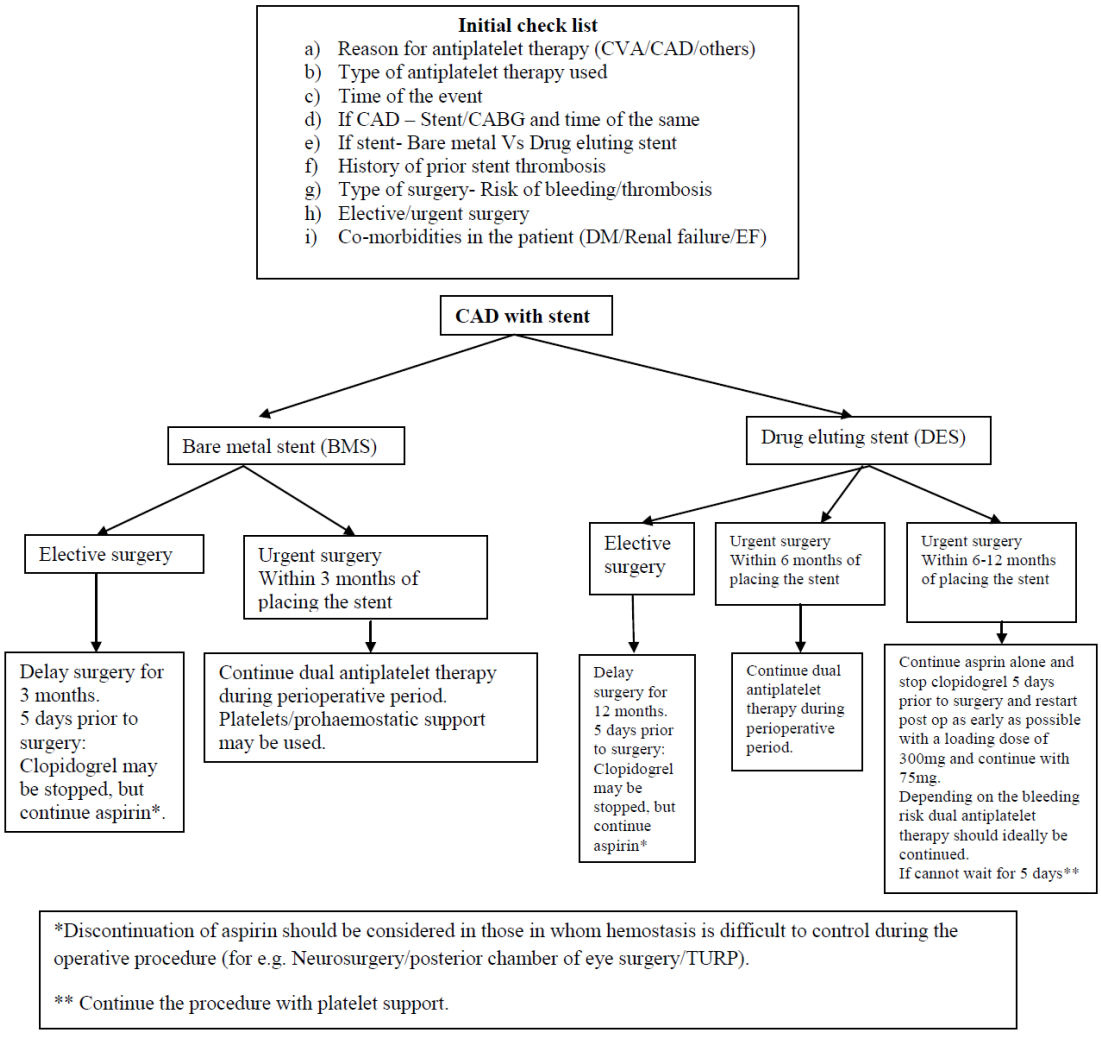
1) Dual anti-platelet therapy (DAPT)

In patients with ACS who have had a PCI with the placement of a stent (drug eluting or bare metal) elective surgery should be delayed and DAPT continued for 12 months with a drug eluting and 4-6 weeks for a bare meta stent. The risk of in-stent thrombosis is highest for bare metal stents in the first 6 weeks and for drug eluting stents within the first 6 months. In the 5% of patients undergoing PCI and stent placement that present for non-cardiac surgery, aspirin should be continued and the P2Y12 anti-platelet agent discontinued if the procedure involves areas such as intra-cranial, intra-spinal, intra-ocular, pericardial and neuraxial locations. If appropriate, the patient should be bridged with LMWH or unfractionated heparin.

Guidelines for patients with bare metal or drug eluting stents on DAPT who present for elective and emergency surgery are illustrated below (*figure 2*).

Figure 2

Decision making algorithm in patients undergoing non-cardiac surgery[6, 8, 9, 22]



*Discontinuation of aspirin should be considered in those in whom hemostasis is difficult to control during the operative procedure (for e.g. Neurosurgery/posterior chamber of eye surgery/TURP).

** Continue the procedure with platelet support.

Management of bleeding

Although the new anti-platelet agents generally have short half-lives, their effects are long lasting (i.e. the life if the platelet). As no reversal agents exist, the mainstay of treatment is to institute general resuscitative measures as required, withdraw anti-platelet medication and, in extreme circumstances, transfuse platelet concentrates, blood products and FFP.

2) Anticoagulants

Similar principles regarding risk of thrombosis against that of bleeding apply to the novel anticoagulants. Although no antidotes exist clinically at present, these drugs generally have a rapid onset and offset so should be easier to manage. The necessity for bridging patients on these drugs prior to surgery can be defined by scoring patients on the CHADS2 and CHA2DS2-VASc scales that assess thrombosis risk in patients with atrial fibrillation.

C	CHF / LV dysfunction	1
H	Hypertension	1
A	Age > 75	2
D	Diabetes mellitus	1
S	Stroke / TIA / Embolism	2
V	Vascular disease	1
A	Age 65-74	1
Sc	Sex-Female	1

A patient with a score of 5 or greater, (maximum 9), stroke or TIA within three months, cancer or a prosthetic valve is felt to be at high risk of stroke, approximately 12-18 per 100 patients per year. When $CHADS_2$ alone is used, a score greater than 2 is associated with increased risk of stroke.

A patient with a CHADS2 score of 2 and above or a CHA2DS2-VASc score of 5 and above should be bridged with therapeutic LMWH. The oral anticoagulant should be stopped 5 days prior to surgery with therapeutic LMWH started three days prior to surgery. A prophylactic risk surgery in patients with normal renal function, the anticoagulant can be stopped 2 days before surgery with the administration f a prophylactic dose of LMWH the evening prior to surgery. In patients with impaired renal function, (eGFR <80mls/min), treat as for major surgery.

Management of bleeding (*table 5*)

Table 5

Multidisciplinary approach

Withhold drug-maybe sufficient for mild bleeding

General haemostatic and resuscitation measures: large bore iv access, bloods and clotting assays, mechanical pressure, fluid and / or blood administration

Mild: as above plus TXA 15-25mg/kg orally

Moderate - severe: as general plus FFP 20 ml/kg + platelet transfusion (if < 70 $x10^9$/L), TXA 15mg/kg iv, consider PCC, e.g. Octaplex® 30iu/kg[(a)]

Life-threatening: as above plus PCC; consider charcoal (if within 2 hours of drug administration) / CVVHF for dabigatran [26]

(a) Octaplex ® is a second generation human prothrombin complex concentrate.

Management of patients presenting for emergency surgery

(see above for patients on DAPT).

Emergency surgery should be started at least 2 half-lives after the last dose of anticoagulant if time and circumstance permits (>12 hours for dabigatran and >24 hours for rivaroxaban and apixaban). If surgery must happen sooner than suggested above for each drug, treatment should follow the guidelines set out in *table 5* above.

Timing of neuraxial block

Clopidogrel and prasugel bind irreversibly to the P2Y12 receptor and thus has antiplatelet effects that last for the life of the platelet. Clopidogrel and prasugel should be stopped for at least 7 days before neuraxial anaesthesia. Clopidogrel may be re-started 2 hours and prasugel 6 hours after epidural catheter removal.

Table 6: Timing of Central Neuraxial Block in patients taking novel anticoagulants

Therapy	Time from dose to insertion / removal	Time from insertion / removal to next dose
Dabigatran	2 days if Creatinine Clearance (CrCl) > 50 mls/min, 5 days if CrCl < 50 mls/min[21]	The first dose of dabigatran may be given two hours following the removal of an epidural catheter, but concomitant use is contraindicated
Rivaroxaban	22-26 hours	6 hours (24 hours if traumatic puncture)
Apixaban	26-30 hours	4-6 hour
Fondaparinux	36-42 hours	6-12 hours

References:

1) Smith F, Telford R. ATOTW-Novel anti-platelet agents and anti-coagulants, 22/09/2014
2) Arora H. New Anticoagulants and Emerging Trends in Regional Anesthesia. *Anesthesiologynews.com*
3) Thuy Anh Nguyen, Jean Diodati, Chantal Pharand. Resistance to Clopidogrel: A Review of the Evidence. JACC Vol. 45, No. 8, April 19, 2005: 1157-64
4) Xi-Ling Jiang, Snehal Samant, Lawrence Lesko *et al.* Clinical Pharmacokinetics and Pharmacodynamics of Clopidogrel. Clin Pharmacokinet (2015) 54: 147-166
5) Guidelines for Neuraxial Anesthesia and Anticoagulation. Intermountain Guidelines. https://kr.ihc.com/ext/Dcmnt?ncid=520499512&tfrm=default
6) Frank Peacock, Zubaid Rafique, Adam Singer. Direct-Acting Oral Anticoagulants: Practical Considerations for Emergency Medicine Physicians. Emergency Medicine International, Volume 2016.
7) Joseph John, Sandeep Chopra. "Management of Patients on Antiplatelet Drugs Undergoing Surgery" To Stop OR Not TO Stop. Medicine Update 2012: Vol 22

The Bohr and Haldane Effects

Dr Kobus Bergh

UCT Dept of Anaesthesia & Perioperative Medicine

The Bohr and Haldane effects have got to do with the loading of oxygen to the haemoglobin molecule and the unloading of oxygen from the haemoglobin molecule [6]. Just in case you were wondering, both these effects got their names from the people who described them.

I will first discuss some essential theory that is needed to understand these effects, before exploring them in detail.

Essential knowledge needed to understand the Bohr and Haldane effects

Some essential facts about haemoglobin [1,2,7]

Haemoglobin is the red, O_2 carrying pigment in the red blood cells (erythrocytes) of vertebrates. There are about 200 – 300 million haemoglobin molecules in each red blood cell.

Haemoglobin is a protein and this protein is made up of 4 subunits (Fig. 1). Each subunit contains a *haem moiety* attached to a polypeptide chain. (Definition of moiety = a distinct part of a large molecule). The *haem moiety* is a complex made up of a porphyrin and a central iron atom in the ferrous state (Fe^{2+}). The polypeptides are referred to collectively as the *globin portion* of the haemoglobin molecule. Some finer detail to take note of is that the *haem moiety* is attached (at a constant distance) to a histidine group on the *globin portion*, and this forms one subunit of haemoglobin.

There are 2 pairs of polypeptides in each haemoglobin molecule. In normal adult human haemoglobin (HbA), the 2 types of polypeptides are called ***α chains***, (each α chain contains 141 amino acid residues), and ***β chains***, (each β chain contains 146 amino acid residues).

Each of the 4 iron atoms can bind reversibly to 1 molecule of oxygen, and therefore each haemoglobin molecule can bind to 4 oxygen molecules. Bear in mind that the iron stays in the ferrous state, so that the reaction is an oxygenation, not an oxidation! HbA can have its ferrous ion (Fe^{2+}) oxidized to the ferric form (Fe^{3+}) by drugs and chemicals such as prilocaine, nitrates, nitrites, sulfonamides, and acetanilid. Deficiency of the enzyme methaemoglobin reductase within the red blood cell whose job it is to convert Fe^{3+} to Fe^{2+} may also cause this. When the iron atom is in its ferric form it is known as methaemoglobin and is unable to carry oxygen. [1,7]

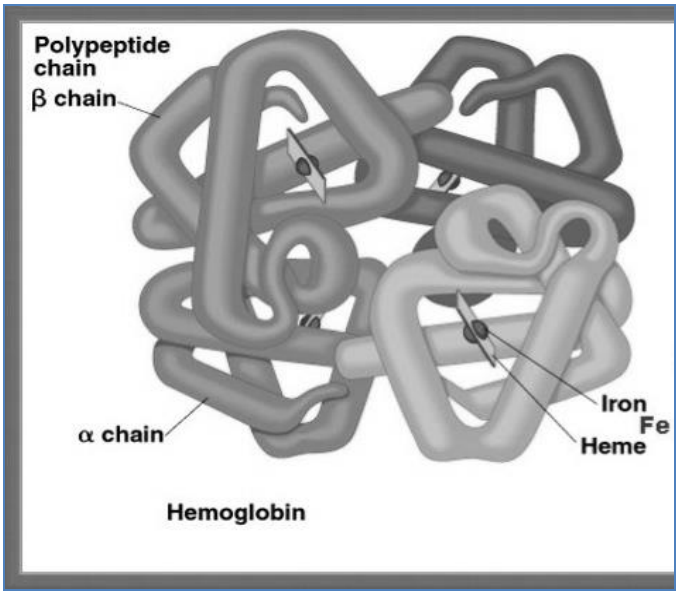


Fig.1 Haemoglobin A

Image taken from: https://bhavanajagat.com/2013/10/26/wholedude-wholedesigner-red-blood-cell

Oxygen transport in the blood [1,7]

It is important to appreciate that oxygen is carried in blood in two forms: *dissolved* and *combined with haemoglobin (Hb).*

Dissolved oxygen

- Like all other gasses, oxygen obeys Henry's Law, which states that the gas dissolved in a liquid is proportional to its partial pressure.
- For each mmHg of PO_2, there is 0,003 ml O_2 dissolved per 100 ml of blood at 37°C.
- Therefore arterial blood with a PO_2 of 100 mmHg (13,3 kPa) contains 0,3 ml oxygen per 100ml.
- It is clear that this way of transporting oxygen is inadequate and that an additional method for transporting oxygen is required…!

Combined with haemoglobin

- I have discussed most of this already under Haemoglobin.
- Taken note that oxygen bound to Hb does not contribute directly to the PO_2 of the blood, only dissolved oxygen contributes to PO_2.
- At normal atmospheric pressure, 98% of oxygen in blood is bound by Hb.

The reaction of haemoglobin and oxygen [2,3,7]

The quaternary structure of haemoglobin determines its affinity for O_2. The change in Hb from the fully oxygenated state to its deoxygenated state is accompanied by a conformational change in the molecule. The oxygenated form is the *R (relaxed) state*, while the deoxygenated form is the *T (tense) state*.

In de-oxyhaemoglobin, the globin units are tightly bound in a *tense (T) configuration*, which reduces the affinity of the molecule for O_2. When O_2 is first bound, the bonds holding the globin units are released, producing a *relaxed (R) configuration*, which exposes more O_2 binding sites. The net result is a 500-fold increase in O_2 affinity! In the tissues these reactions are reversed and O_2 is released.

The oxygen-haemoglobin dissociation curve (which will be extensively discussed in another lecture) is a curve that plots oxygen saturation of haemoglobin against PaO_2. It has its characteristic sigmoid shape due to the T-R interconversion. Combination of the first haem in the haemoglobin molecule with O_2 increases the affinity of the second haem for O_2, and oxygenation of the second, increases the affinity of the third and oxygenation of the third increases the affinity of the fourth! Therefore the affinity of haemoglobin for the fourth O_2 molecule is many times that for the first.

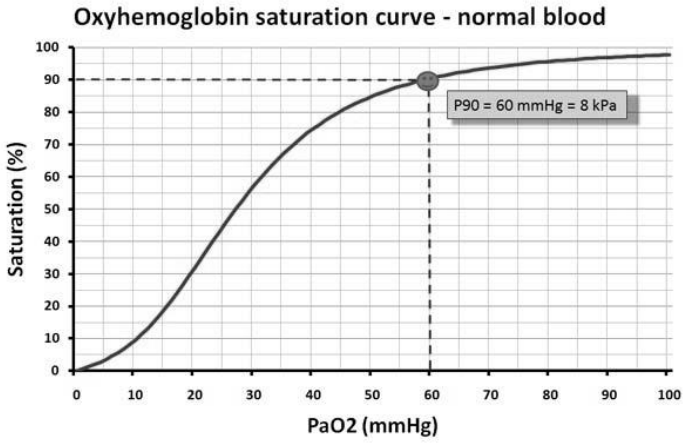


Fig. 2 The Oxygen-haemoglobin dissociation curve

- *Image taken from: http://engineering.stackexchange.com/questions/3058/how-to-generate-a-synthetic-infrared-and-red-led-signal-for-spo2-calculation*

Factors affecting the affinity of haemoglobin for oxygen [2, 7]

Five important conditions affect the oxygen-haemoglobin dissociation curve:

- the **pH**
- the **hydrogen ion** concentration
- the amount of **CO_2**
- the **temperature**
- the concentration of **2,3 diphosphoglycerate (2,3 DPG)**.

A rise in temperature or a fall in pH shifts the curve to the right (Fig 3.7). A rightward shift means more unloading of oxygen at a given PO_2 in a tissue capillary. A simple way to remember these shifts is that an exercising muscle is acidic, hypercarbic, and hot, and it benefits from increased unloading of oxygen from its capillaries. When the curve is shifted to the right, a higher PO_2 is required for haemoglobin to bind a given amount of O_2. On the other hand, a fall in temperature or a rise in pH shifts the curve to the left, and a lower PO_2 is required to bind a given amount of O_2. A convenient index of such shifts is the P_{50}. The P_{50} is the PO_2 at which haemoglobin is half saturated with oxygen. The higher the P_{50}, the lower the affinity of haemoglobin for oxygen.

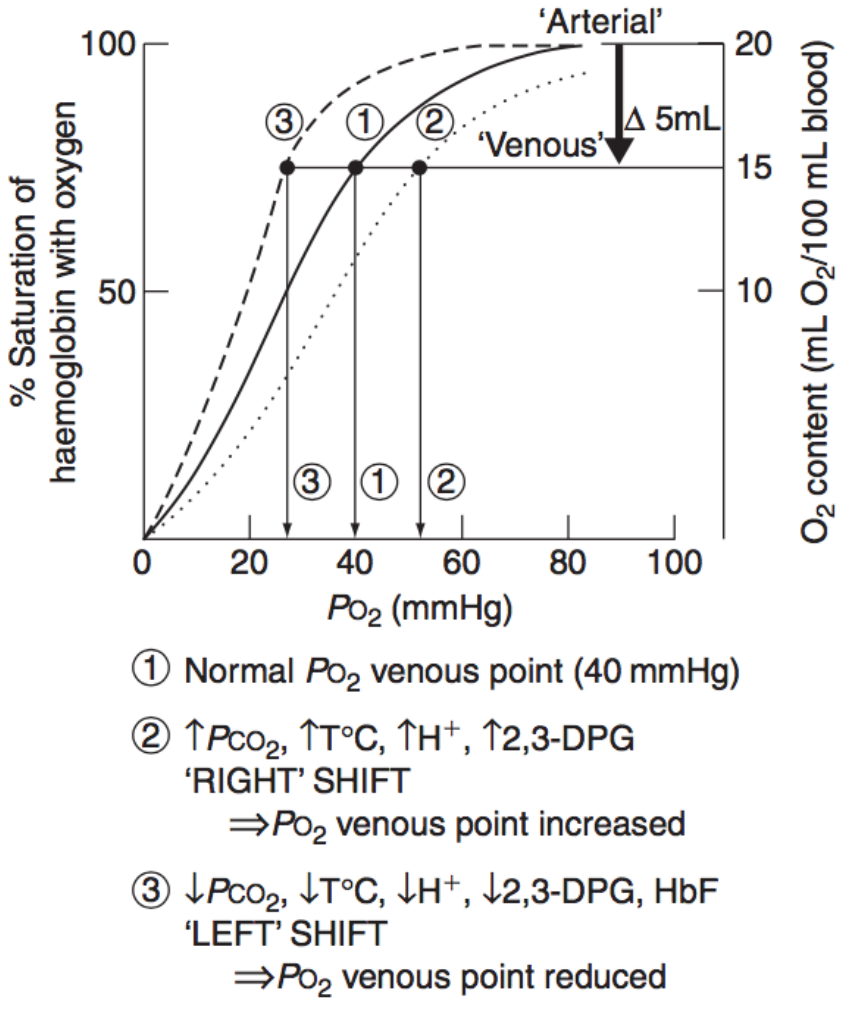


Figure 3.7 *Graphical representation of the Bohr effect. 2,3-DPG, 2,3-diphosphoglycerate.*

Taken from Power I, Kam P. Chapter 3: Respiratory physiology. Principles of Physiology for the Anaesthetist, second edition. (p. 84). Hodder Arnold

Carbon dioxide transport in the blood [2, 7]

It is important to know that CO_2 is carried in blood in the following three forms:

1. Dissolved (in plasma and RBC)
2. Bicarbonate (in plasma and RBC)
3. In combination with proteins as carbamino compounds (in plasma and RBC)

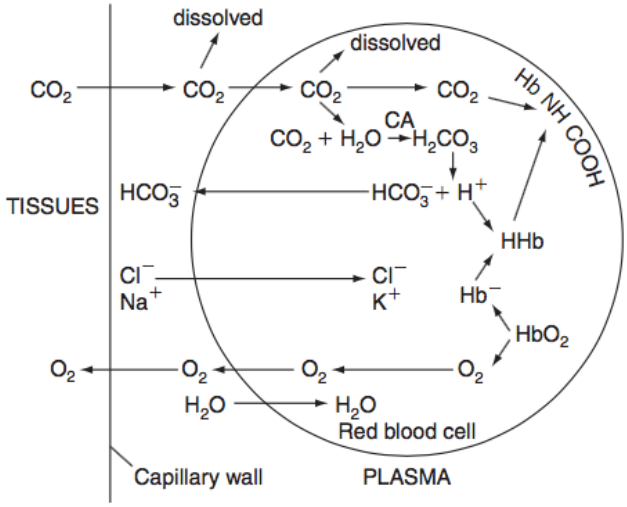


Figure 3.9 *Diagrammatic representation of the uptake of CO_2 and liberation of O_2 in systemic capillaries. CA, carbonic anhydrase. (Reproduced, with permission, from* Respiratory Physiology – the essentials, *5th Edition. John B West, 1995, Williams and Wilkins.)*

Taken from Power I, Kam P. Chapter 3: Respiratory physiology. Principles of Physiology for the Anaesthetist, second edition. (p.86). Hodder Arnold

1. Dissolved CO_2

Like O_2, CO_2 also obeys Henry's Law. CO_2 is however about 20 times more soluble than O_2 in simple solution at equal partial pressures. As a result, dissolved CO_2 plays a significant role in its carriage in blood. About 10% of the gas that is evolved into the lung from the blood is in the dissolved form.

2. Bicarbonate

Bicarbonate is formed in the blood by the following sequence:

$$CO_2 + H_2O \underset{\text{CA (carbonic anhydrase)}}{\leftrightarrows} H_2CO_3 \rightleftarrows H^+ + HCO_3^-$$

The CO_2 that diffuses into red blood cells is rapidly hydrated to H_2CO_3. The fist reaction is very slow in plasma but is fast within the red blood cell, because of the presence of the enzyme carbonic anhydrase (CA) in the red blood cell. The second reaction, which is the ionic dissociation of carbonic acid to form H^+ and HCO_3^- is fast without an enzyme!

What happens to the formed HCO_3^-?
Well, this is the ideal place to discuss the Chloride Shift…

The Chloride Shift (Hamburger phenomenon)
The rise in the HCO_3^- content of red cells is much greater than that in plasma as the blood passes through the capillaries, and about 80% of the HCO_3^- formed in the red blood cells diffuses out into the plasma. H^+ cannot easily diffuse out because the red cell's membrane is relatively impermeable to

cations. Therefore to maintain electrical neutrality, Cl^- ions move into the red blood cell from the plasma. This is the so-called chloride shift, also known as the Hamburger phenomenon, named after Hartog Jakob Hamburger.The chloride shift is responsible for the fact that the chloride content of the red cells in venous blood is significantly greater than in arterial blood. The chloride shift occurs rapidly and is essentially complete in 1 second. Note that for each CO_2 molecule added to a red cell, there is an increase of one osmotically active particle, either an HCO_3^- or a Cl^- in the red cell. Consequently the red cells take up water and increase in size! When the cells pass through the lung again they shrink a little!

What happens to the formed H^+?
Some of the H^+ ions bind to de-oxyhaemoglobin, which is illustrated below.

$$H^+ + HbO_2 \leftrightarrows H^+ - Hb + O_2$$

This occurs because de-oxyhaemoglobin is less acidic and is therefore a better proton acceptor than the oxygenated form. The presence of de-oxyhaemoglobin in the peripheral blood helps with the loading of CO_2, while the oxygenation that occurs in the pulmonary capillary assists in the unloading. The fact that deoxygenation of the blood increases its ability to carry CO_2 is known as the Haldane effect. (I will explain the Haldane effect in much more detail later!)

3. Carbamino compounds

Carbamino compounds are formed by the combination of CO_2 with terminal amine groups in blood proteins. The most important protein is the globin of haemoglobin, and is illustrated by the following reaction:

$$Hb\text{-}NH_2 + CO_2 \rightleftarrows Hb\text{-}NH\text{-}COOH \text{ (This is carbamino-haemoglobin)}$$

This reaction occurs rapidly without an enzyme, and de-oxyhaemoglobin ($Hb\text{-}NH_2$) can bind more CO_2 to form carbamino-haemoglobin (Hb-NH-COOH) than oxyhaemoglobin (HbO_2). I hope that it is clear to you that the unloading of O_2 in peripheral capillaries, facilitates the loading of CO_2, while oxygenation in the lungs has the opposite effect.

A last word on CO_2 carriage…
Note that the greatest bulk of the CO_2 is in the form of bicarbonate (80 - 90%). The amount of dissolved CO_2 is small (5 - 10%), as well as the amount of CO_2 carried as carbamino-haemoglobin (5 - 10%).

Summary of Carbon Dioxide Transport [2]

In Plasma	*In Red Blood Cells*
1. Dissolved 2. Formation of carbamino compounds with plasma proteins 3. Hydration, H^+ buffered, HCO_3^- in plasma	1. Dissolved 2. Formation of carbamino-Hb 3. Hydration, H^+ buffered, large% of HCO_3^- enters the plasma 4. Cl^- shifts into cells

Of the approximately 49 ml of CO_2 in each 100ml of arterial blood, 2,6 ml is dissolved, 2,6 ml is in carbamino compounds, and 43,8 ml is in HCO_3^-. In the tissues 3,7 ml of CO_2 per 100ml of blood is added; 0,4 ml stays in solution, 0,8 ml forms carbamino compounds, and 2,5 ml forms HCO_3^-. The pH of the blood drops from 7,40 to 7,36. In the lungs, the process is reversed, and 3,7 ml of CO_2 is discharged into the alveoli. In this fashion, 200 ml of CO_2 per minute at rest and much larger amounts during exercise are transported from the tissues to the lungs and excreted.

CO_2 Dissociation curve [1,7]

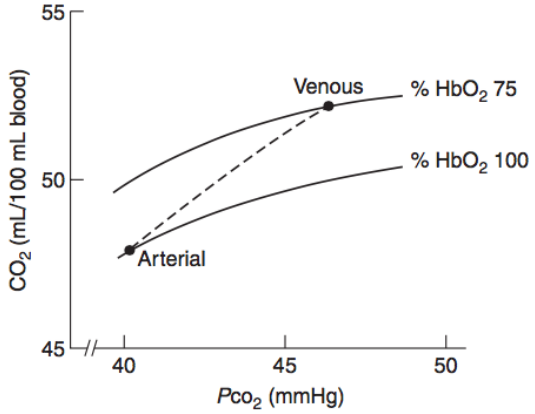


Figure 3.8 *The carbon dioxide–blood dissociation curve.*

Taken from Power I, Kam P. Chapter 3: Respiratory physiology. Principles of Physiology for the Anaesthetist, second edition. (p.86). Hodder Arnold

In contrast to the dissociation of oxygen from haemoglobin, the dissociation of CO_2 from blood is directly related to the PCO_2 and therefore the dissociation curve for CO_2 is linear. Note also that the lower the saturation of Hb with O_2, the larger the CO_2 concentration for a given PCO_2, This is the Haldane effect, and a detailed explanation will follow later.

The Bohr effect[1,2,4,6]

The decrease in oxygen affinity of haemoglobin when the pH of blood falls is called the Bohr effect and is related to the fact that de-oxygenated haemoglobin binds H^+ more actively than does oxyhaemoglobin.

The pH of blood falls as its CO_2 content increases, so that when the PCO_2 rises, the oxygen haemoglobin dissociation curve shifts to the right and the P_{50} rises. Note that most of the unsaturation of haemoglobin that occurs in the tissues is secondary to the decline in the PO_2, but an extra 1-2% of haemoglobin's unsaturation is due to the rise in PCO_2 and the consequent shift of the dissociation curve to the right (Fig. 3.7)

The Bohr effect takes place in tissues that are metabolically active, for example in the quadriceps muscles when you are walking! As I have mentioned earlier this magnificent piece of physiology helps with the unloading of oxygen from the haemoglobin molecule and it does so proportionally to the metabolic activity of the tissue. As more metabolism takes place, the carbon dioxide partial pressure increases and will cause larger reductions in local pH and in turn will allow for greater oxygen unloading. This is especially true in exercising skeletal muscles which may also release lactic acid that further reduces local blood pH and therefore enhances the Bohr Effect.

Now its time for a simplified explanation [6], some parts of this explanation have been discussed under the carriage of CO_2 already but repetition is good!
Imagine an individual quadriceps muscle cell in which CO_2 will be produced due to the Krebs cycle in the mitochondria. The produced CO_2 will diffuse out of the muscle cell down a partial pressure gradient into the interstitium, across the capillary wall, and into the plasma. Some CO_2 will dissolve in the plasma, and some will dissolve in the erythrocyte.
From our essential knowledge discussion, remember that approximately 10% of the CO_2 will directly dissolve in the plasma, and 90% will diffuse into the red blood cell (RBC). You may ask, what will happen to the 90% that diffuses into the RBC? Well 10% of the 90% of CO_2 that diffuses into the RBC will bind directly to haemoglobin to form carbamino haemoglobin and the remaining 80% will combine with H_2O to form H_2CO_3 (carbonic acid) with the help of the enzyme *carbonic anhydrase (CA)*.

$$CO_2 + H_2O \rightleftarrows H_2CO_3$$

(↑ carbonic anhydrase (present in RBC's and endothelium, not plasma))

H_2CO_3 will spontaneously dissociates into H^+ (hydrogen ion) and HCO_3^- (bicarbonate)

What happens with this formed HCO_3^- and H^+?

The HCO_3^-:
The HCO_3^- will leave the RBC and remember that 80% of the CO_2 produced by the quadriceps muscle will be transported to the lungs as bicarbonate in the plasma.
When HCO_3^- leaves the cell, Cl^- will enter the cell to maintain electrical neutrality. This is the *chloride shift*, which we have discussed earlier.

The H^+:
The H^+ will combine with the haemoglobin molecule. The H^+ ions bind to the α-amino and imidazole groups of haemoglobin and alter the allosteric conformation of the haemoglobin, which reduces the affinity of oxygen to haem [1]. Remember that this haemoglobin molecule is saturated with 4 oxygen molecules because it is coming from the lungs.
The H^+ will protonates the histidine residue in the haemoglobin molecule and this will cause stabilization of the T state. The T state of haemoglobin is not favourable for oxygen to be bound to haemoglobin and oxygen will be released for usage by the tissues.

The Haldane effect[1,2,7]

The increased capacity of deoxygenated haemoglobin to carry CO_2 is referred to as the Haldane effect.

De-oxygenated haemoglobin binds more H^+ than oxyhaemoglobin and de-oxygenated haemoglobin forms carbamino compounds more readily than oxyhaemoglobin.

Let's dissect and explore above statement:

De-oxygenated haemoglobin forms carbamino compounds more readily than oxyhaemoglobin
Explanation:
The attachment of oxygen to haem reduces the capacity of haemoglobin to carry CO_2. The reason for this is that O_2 increases the ionization of nitrogen groups, which reduces the capacity of the globin chain to carry CO_2 as carbamino compounds. De-oxyhaemoglobin can carry more CO_2 in the form of carbamino compounds, which account for about one third of the arterial venous difference of CO_2 carried in blood.

De-oxygenated haemoglobin binds more H^+ than oxyhaemoglobin
Explanation:
De-oxyhaemoglobin is more basic (due to the large number of imidazole groups present in the histidine moieties) and therefore has an increased capacity to mop up the H^+ ions produced when carbonic acid dissociates and so has an increased buffering capacity for CO_2.

Consequently, venous blood carries more CO_2 than arterial blood, and CO_2 uptake is facilitated in the tissues and CO_2 release is facilitated in the lungs.

A simplified way to explain the Haldane effect [6]

Let me set the scene:
Imagine a RBC that is returning from the hard working quadriceps muscle and is now arriving in a pulmonary capillary adjacent to an alveolus. Remember that the haemoglobin molecule in this RBC has given off oxygen to the quadriceps muscle. The P_AO_2 of the alveolus = 90 – 100 mmHg (12 – 13,3 kPa). In the plasma of this pulmonary capillary is HCO_3^-.This HCO_3^- will diffuse into the RBC, and as this happens, Cl^- will diffuse out of the cell (remember that at the quadriceps muscle, HCO_3^- diffused out of the RBC, and Cl^- diffuses into the RBC)
The haemoglobin in this RBC (de-oxyhaemoglobin) got a H^+ attached to it as well as CO_2 .
Now that the scene is set, let's see what happens:

1. Oxygen diffuses down its partial pressure gradient from the alveolus into the RBC.
2. As oxygen enters the RBC, it combines with haemoglobin.
3. When oxygen binds to haemoglobin, the H^+ and CO_2 are released from the haemoglobin molecule.
4. What happens to the released H^+?
    a. Well, H^+ will combine with HCO_3^- to form H_2CO_3 (carbonic acid)
    b. H_2CO_3 (carbonic acid) will dissociate with the help of CA (carbonic anhydrase) into $CO_2 + H_2O$
    c. The CO_2 formed from the dissociation of H_2CO_3 will diffuse out of the cell and will diffuse down a partial pressure gradient into the alveolus from where it will get exhaled!
5. What happens to the CO_2 that is released from the haemoglobin?
    a. This CO_2 will also diffuse out of the cell and will also diffuse down a partial pressure gradient into the alveolus from where it will get exhaled

The Double Bohr effect [1,8]

The double Bohr effect got to do with the exchange of oxygen and carbon dioxide between the mother and fetus. Let me set the scene to explain the double Bohr effect:

Think about the mother and fetus and think about the total amount of oxygen that goes from the mother to the fetus. Let's quickly revise some important anatomy that will help us to understand.

Umbilical cord:
Remember that there are 2 umbilical arteries and 1 umbilical vein in the umbilical cord. The umbilical arteries carry de-oxygenated blood from the fetus to the placenta and the umbilical vein carries oxygenated blood from the placenta to the fetus.

Placental structure:
The umbilical arteries branches into the chorionic plate. On the maternal side of the chorionic plate is a pool of blood. The chorionic plate got little extensions that dip into this pool of blood. In these extensions are fetal capillaries, which are tiny extensions of the umbilical vein and umbilical arteries. The mother's uterine arteries also open into the pool of blood, supplying oxygenated blood and the mother's uterine veins drain deoxygenated blood to the mother's lungs. Behind the pool of blood on the mother's side is the thick muscular uterine wall.

Let's look at the oxygen content on the mother's side and the oxygen content on the side of the fetus:

	PO_2	**HbA saturation**	**HbF saturation**
Uterine artery (mother)	100 mmHg	98%	
Uterine vein (mother)	40 mmHg	75%	
Umbilical artery (fetus)	18 mmHg		45%
Umbilical vein (fetus)	28 mmHg		70%

Have a look now at fig 14.12. Look at the differences between the 4 drawn curves.

- The **umbilical artery** has a higher carbon dioxide content and lower pH (and higher hydrogen ion concentration) than the **umbilical vein**.
- The difference in oxygen saturation between the umbilical artery and umbilical vein lines is called the Bohr effect.
- The **uterine vein** has a higher carbon dioxide content and lower pH (and higher hydrogen ion concentration) than the **uterine artery.**
- Similarly, the difference in oxygen saturation between the uterine artery and uterine vein is also called the Bohr effect.

Remember from our discussion earlier, the Bohr effect happens when carbon dioxide and hydrogen ions makes oxygen 'fall off' haemoglobin or haemoglobin doesn't bind oxygen well in the presence of carbon dioxide and hydrogen ions.

Let's discuss another obvious difference between the 2 curves of the fetus and the 2 curves of the mother.

- Appreciate that the fetal curves (umbilical artery curve and umbilical vein curve) are pushed to the left, this is because the fetus got haemoglobin F which has a higher affinity for oxygen compared to haemoglobin A.

Finally, let's look at the Bohr effects:

Bohr effect on the fetal side:

- The Bohr effect takes place when the release of carbon dioxide from the fetal blood inside the chorion enhances the uptake of oxygen. To illustrate this on the diagram, it is the vertical difference between the umbilical artery curve and the umbilical vein curve.

Bohr effect on the maternal side:

- Inside the pool of blood, the carbon dioxide levels are slowly rising, and the Bohr effect causes the release of oxygen molecules from the maternal haemoglobin, which causes a right shift of the curve. Again it can be illustrated on the diagram by the vertical distance between the uterine artery and uterine vein curve.

Because both these Bohr effects are happening in the placenta we call it the double Bohr effect. The double Bohr effect refers to the 4 lines illustrated in the diagram. All of this is happening in the placenta at the same time!

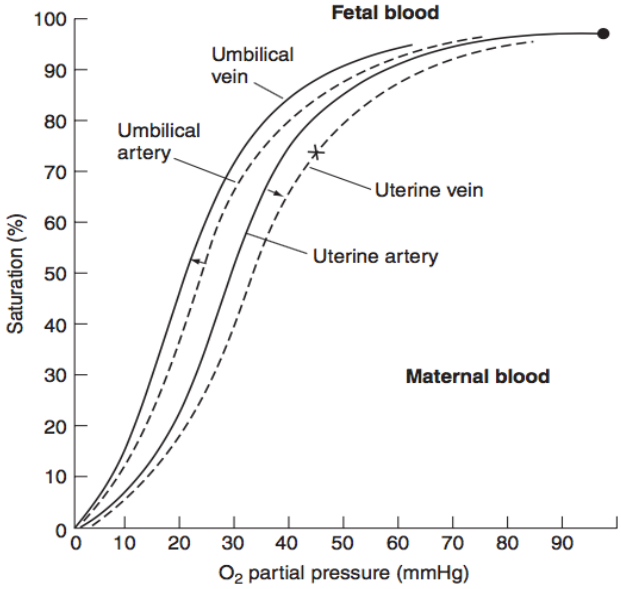


Figure 14.12 *Transport of oxygen from the mother to the fetus: the double Bohr effect.*

Taken from Power I, Kam P. Chapter 3: Respiratory physiology. Principles of Physiology for the Anaesthetist, second edition. (p.86). Hodder Arnold

References

1. Power I, Kam P. Chapter 3: Respiratory physiology. *Principles of Physiology for the Anaesthetist,* second edition. (pp. 81 – 86). Hodder Arnold.
2. Ganong WF. Chapter 35: Gas transport between the lungs and the tissues. *Review of Medical Physiology*, twenty-first edition. (pp. 669 – 674). Lange.
3. Yentis S, Hirsch N, Smith G. *Anaesthesia and Intensive Care A-Z*, fourth edition. Churchill Livingstone Elsevier.
4. Retrieved from http://www.pathwaymedicine.org/bohr-effect
5. Retrieved from https://www.khanacademy.org/science/health-and-medicine/advanced-hematologic-system/hematologic-system-introduction/v/bohr-effect-vs-haldane-effect
6. Retrieved from https://youtu.be/PM6maiHEnIM
7. JB West. Chapter 6: Gas transport by the blood. *Respiratory Physiology, The Essentials, seventh edition*. (pp. 75 – 83).
8. Retrieved from https://www.khanacademy.org/science/health-and-medicine/circulatory-system/fetal-circulation/v/double-bohr-effect

Uterotonic and Tocolytic drugs
What should the anaesthesiologist know?

Prof. Rob Dyer

UCT Dept of Anaesthesia & Perioperative Medicine

Many drugs are administered with the specific purpose of influencing uterine tone. The two main classes and indications are uterotonic agents for the prevention and treatment of postpartum haemorrhage (PPH), and tocolytic drugs for preterm labour. These agents have a narrow therapeutic range in terms of maternal and fetal morbidity. The exact dose, route and rate of administration are therefore important, as well as a detailed knowledge of the pharmacology.[1]

A: Uterotonic agents

Introduction

Every year 166,000 women die of obstetric haemorrhage, and more than 50% of these deaths occur in sub-Saharan Africa. Uterine atony is the commonest cause of severe PPH. Uterotonic drugs are thus an essential pharmacological intervention by anaesthetists during caesarean section (CS), in order to diminish the risk of PPH and improve maternal safety.

Oxytocin

The nonapeptide oxytocin was discovered by Sir Henry Dale and was the first polypeptide hormone synthesised, by Du Vigneaud, in 1953. The peptide binds to a G-protein on the surface of the uterine myocyte, resulting in the generation of diacylglycerol (DAG) and inositol tri-phosphate (IP3) via the action of phospholipase C on phosphatidyl inositol bisphosphate (Figure 1). DAG stimulates prostaglandin synthesis, and IP3 stimulates the release of calcium from the sarcoplasmic reticulum. The resulting Ca^{2+}-Calmodulin complex activates myosin light chain kinase and results in phosphorylation of myosin and further initiation of the actin-myosin ATPase, and hence contraction. Oxytocin also activates COX-2 via a further G-protein interaction, and in so doing stimulates prostaglandin synthesis.

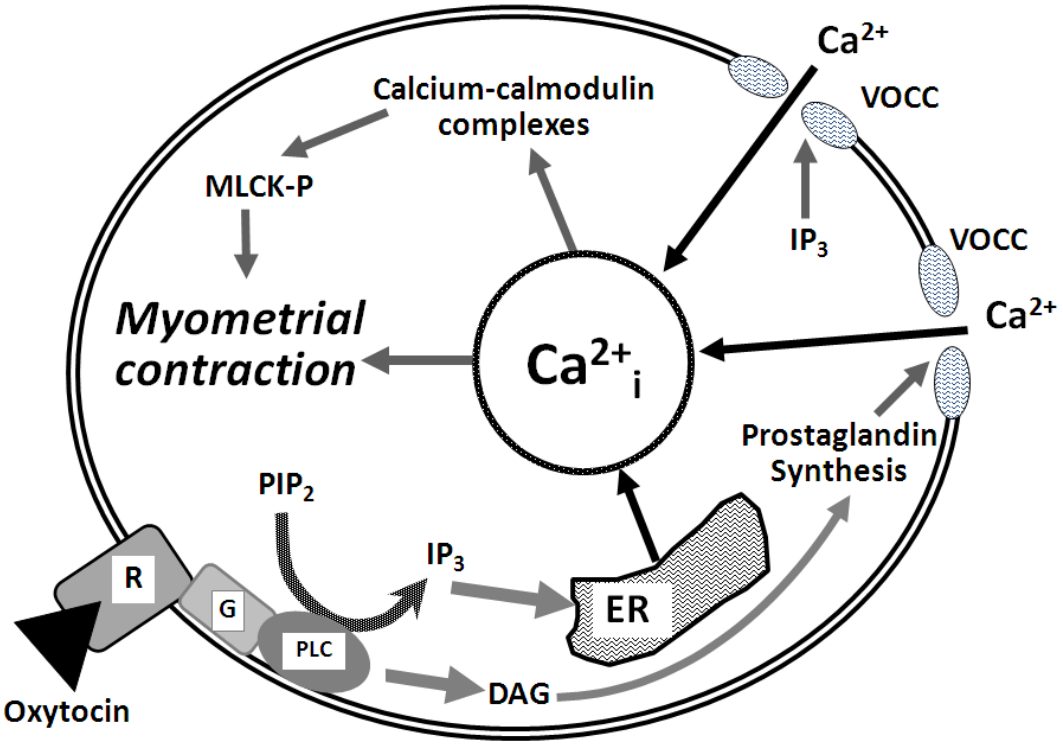


Figure 1: Mechanism of action of oxytocin

The concentration of myometrial receptors as well as myometrial gap junctions increase as gestation advances, thus increasing sensitivity to oxytocin. Oxytocin has numerous physiological effects. Most importantly, it causes contraction, followed by relaxation of the uterus, and at pharmacological doses can cause an increased frequency and incomplete relaxation of the uterine musculature. It also has a role in cardiovascular regulation, in sexual and maternal behaviour, and in memory and the regulation of food and drink intake. This agent remains the first line uterotonic during CS.

Major maternal adverse effects are cardiovascular (hypotension, myocardial ischaemia and arrhythmias), nausea, vomiting, headache and flushing. The cardiovascular effects are complex. Hypotension is predominantly caused by transient relaxation of vascular smooth muscle cells, probably via calcium dependent stimulation of the nitric oxide pathway. Oxytocin also causes the release of atrial and brain natriuretic peptide. In animal studies, oxytocin may have negative inotropic effects, but in human tissue, this effect appears to be restricted to the influence of its commercial preservative chlorbutanol on atrial myocytes in vitro. Due to structural similarities with vasopressin, overdose of oxytocin may cause water retention, hyponatraemia, seizures and coma.

Recently, studies using beat by beat pulse wave form monitors transthoracic bioimpedance, have elucidated a clinical picture of peripheral vasodilatation, hypotension, and increased cardiac output mediated by an increase in heart rate and stroke volume. Pulmonary artery pressures are markedly increased and sustained for at least 10 minutes after a bolus of 10 IU during general anaesthesia. One observational study has demonstrated similar effects of 2.5 IU on the systemic vasculature in patients with severe preeclampsia. These effects could be poorly tolerated if ventricular function were abnormal, and in the presence of mitral or aortic stenosis, or hypovolaemia. A fatality was recorded in the Confidential Enquiry into Maternal Deaths of the United Kingdom in the triennium 1997-1999, when oxytocin 10IU was administered during the resuscitation of a hypovolaemic patient with probable high spinal anaesthesia (SA) for CS. In the Report on Confidential Enquiries into Maternal Deaths in South Africa for the triennium 2005-2007, there were 2 deaths in which oxytocin was contributory. In one case a high dose compounded spinal hypotension. In the other, a poorly resuscitated patient undergoing emergency CS received 10 IU of oxytocin and a fatal cardiac arrest ensued.

In contrast to the effects on the systemic vasculature, oxytocin has been shown to cause coronary vasoconstriction in an isolated dog heart model. There may also be considerable ST segment changes following 10 IU oxytocin during SA for CS and in volunteers. A Holter investigation has demonstrated more ST segment depression in healthy women during elective CS under SA, who received a 10 IU- rather than 5 IU bolus of oxytocin. These complications and research findings indicate that a rapid bolus of 10 IU is no longer advised after delivery at CS.

How can the cardiovascular effects of oxytocin be obtunded? A comparison of 2 IU with 5 IU as a bolus showed more minor heart rate and blood pressure changes after 2 IU. The administration of 5 IU of oxytocin by slow infusion has been shown to produce less cardiovascular instability than a bolus of 5 IU. In an observational study, oxytocin was used in incremental doses of 0.1-0.5 IU during CS in parturients with advanced cardiac disease, including cardiomyopathy, congenital and valvular heart disease, with acceptable haemodynamic stability. Co-administration of phenylephrine with oxytocin may obtund the peripheral vascular effects, with some overshoot of the effects of phenylephrine. A recent study showed that administration of phenylephrine prior to bolus oxytocin is relatively ineffective in obtunding hypotension. An interesting animal study has shown that that the effects of phenylephrine on cardiac output are dependent on the position on the Frank-Starling curve.

In view of the multiple side effects of oxytocin, it is desirable to administer the lowest possible effective dose in the most stable manner. The dose and rate of intravenous infusion of oxytocin after delivery during CS remain controversial. There have been only three dose-finding studies. The first study in healthy uncomplicated pregnancies at low risk for uterine atony, showed that the ED90 is 0.35 IU. In a case series of patients with labour arrest, the ED90 was found to be 3.0 IU. A pragmatic recent editorial advocates the "rule of threes". Up to 3 times 3 IU oxytocin are administered at 3 minute intervals no faster than 15 seconds, followed by slow infusion and second line pharmacological options as necessary.[3]

An early investigation showed that when compared with vasopressin, oxytocin is much less active as an antidiuretic when the infusion rate is less than 45 milliunits/minute. This suggests that the rate of infusion for prophylaxis of PPH should be restricted to a lower infusion rate than this, particularly in patients with preeclampsia, who are at higher risk of pulmonary oedema. The current recommended infusion rate for PPH prophylaxis in South Africa is 20 IU/L at 100-150 mL/h. If this cannot be reliably administered, it may be better to give either oxytocin 10 IU IM, repeated at 4 h, or syntometrine IM.[4] For treatment of PPH, the Royal College of Obstetricians and Gynaecologists (RCOG) recommend a

rapid infusion oxytocin 40 IU/500 mL crystalloid at 125 mL/hour, until haemorrhage is controlled. Great care should be exercised in the hypovolaemic patient, and effective resuscitation and vasopressor therapy should accompany oxytocin administration.

In keeping with the mechanism of action of oxytocin, involving G-protein receptor interactions, the phenomenon of receptor desensitisation may influence the effectiveness of the dose given by the anaesthetist at delivery. Definitive laboratory work has shown that there is loss of oxytocin receptors during oxytocin-induced and oxytocin-augmented labour. The concentration of oxytocin receptors decreased more than 3-fold, and oxytocin receptor mRNA concentrations decreased 60- and 300-fold during oxytocin-augmented and oxytocin-induced labour respectively. In view of the fact that repeated doses of oxytocin may become increasingly ineffective, second-line uterotonic agents are still required.

The in vivo half-life of oxytocin is only 5-10 minutes. The newly developed synthetic analogue of oxytocin, **carbetocin**[5], an octapeptide (1-desamino-1-monocarbo-[2-O-methyltyrosine]-oxytocin), has a half-life of 42 minutes. Carbetocin produces contractions of sustained higher frequency and amplitude. The advantage is thus that carbetocin may be used as a single bolus dose. IV carbetocin 8-30 µg causes a tetanic contraction within 2 minutes and lasting 6 minutes, followed by rhythmic contractions for 60 minutes. IM injection 10-70 µg causes tetanic contraction for 11 minutes, followed by rhythmic contractions for 2 hours. Clinical dose-finding studies have used 20-120 µg. The incidence of hypotension and other side effects is similar to that associated with oxytocin, and carbetocin also exhibits receptor desensitisation. The ED90 has been found to be 14.8 µg at elective CS, <20% of the current recommended dose of 100 µg. In cases of labour arrest this increases to 121 µg. Carbetocin is associated with lower risk of PPH than oxytocin, as well as a reduced requirement for additional uterotonics after CS. After vaginal delivery, the mean fall in Hb is lower than when oxytocin/ergometrine is administered. There is no evidence that carbetocin reduces the incidence of severe PPH. A new room temperature stable formulation of carbetocin is under investigation. This could have important application in limited resource environments.

Ergot alkaloids

Ergot, derived from the fungus *Claviceps purpurea*, was the first effective oxytocic drug. The fungus was known as “the noxious pustule in the ear of grain” since 600 BC, owing to epidemics of ergotism, characterised by either central nervous complications or peripheral gangrene. These agents have been used in obstetrics from 1582-1822, when the view on their role changed from “pulvis ad partum” (the powder of birth), to “pulvis ad mortem” (the powder of death), due to the associated tetanic uterine contractions, leading to fetal asphyxia, stillbirth and uterine rupture. Ergometrine is a naturally occurring alkaloid, first isolated in 1932 by Dudley and Chassar Moir. Currently this agent, although appropriately banned from intrapartum use, remains the second line intervention in the absence of contraindications, if uterine atony persists after oxytocin administration during caesarean delivery.

Ergometrine maleate or methylergometrine cause a rapid and sustained contraction of the pregnant and non-pregnant uterus. The half-life of ergometrine is 120 minutes. Little is known about the mechanism of action, which may be via a calcium channel, or an α-receptor in the inner myometrial layer. Ergometrine is also a partial agonist at α-adrenergic, 5HT-1, and dopamine receptors.

Although ergometrine and methylergometrine have the least vasoconstrictor effects of all the ergot alkaloids, the use of ergometrine has been associated with a mean arterial pressure increase of 11% after 0.2 mg IV, and pulmonary wedge pressure and pulmonary artery pressure increases of 30%. There are also reported cases of renal and coronary artery spasm, and there are several cases of myocardial infarction in the literature associated with their use. In some of these cases the use of ergot alkaloids was inappropriate. One patient had familial hypercholesterolaemia and ergometrine precipitated the requirement for stenting of the left anterior descending coronary artery. In another case a fatal myocardial infarction followed the administration of ergometrine to a hypertensive patient with preeclampsia. Despite these rare complications, ergot alkaloids still have an important role as a second line agent at CS when administered with due care, but are contraindicated in preeclampsia.

The RCOG guideline recommends 0.5 mg IV *slowly* for the management of uterine atony, while the World Health Organisation (WHO) stipulates 0.2 mg IV or IM, to be repeated as necessary every 15 minutes to a maximum of 1 mg. The high incidence of nausea and vomiting after the recommended 0.5 mg dose has discouraged its use as a first line agent at CS. **Syntometrine** is a combination preparation seldom used during CS, containing 5 IU oxytocin and 0.5 mg ergometrine. Following IM administration, the time to onset of the uterine response is considerably shorter than after ergometrine alone, and the duration of action is several hours.

Prostaglandins

Like oxytocin, prostaglandins increase intramyometrial calcium concentrations and enhance uterine contraction. Their effects are mediated via G-proteins and the activation of a calcium channel. Side effects after pharmacological administration include fever, diarrhoea, nausea, vomiting and pyrexia. The use of intramyometrial prostaglandin F2α (dinoprost) for atonic PPH was first described by Takagi in 1976. Subsequently, 15-methyl prostaglandin F2α (carboprost) was shown to have an extended half-life, fewer gastrointestinal and vasopressor side-effects, and good uterotonic activity. Since 15-methyl prostaglandin F2α may be associated with bronchospasm, ventilation perfusion mismatch and hypoxaemia, this agent is best used as a last resort therapy and not as prophylaxis. There is very limited experience with intravenous administration. Infusion at 100 μg/minute during early pregnancy has been shown to cause systemic and pulmonary hypertension, in contrast with prostaglandin E2, which is associated with a marked decrease in systemic vascular resistance and hypotension. The recommended dose is 250 μg IM, repeated every 15 minutes to a maximum of 8 doses. Carboprost 500 μg may be administered intramyometrially, but this remains the responsibility of the clinician.

Prostaglandin E1 (misoprostol) is a cheap and widely available oxytocic, which is less effective than oxytocin and ergometrine for the prevention of PPH, and confers no benefit in the management of PPH in women who have received conventional uterotonics, but does however have a role as a first line agent where the former are unavailable (www.misoprostol.org). The sublingual route is probably the most reliable, since misoprostol is a methyl ester and undergoes first pass elimination. Misoprostol is frequently used off-label via the sublingual or rectal route for the management of uterine atony. There have been a limited number of studies of its use during CS. One study concluded that buccal misoprostol may reduce the need for additional uterotonic agents at CS. In a further small randomised comparison, oral misoprostol was concluded be as effective as intravenous oxytocin in reduction of intra-operative blood loss during elective CS under regional anesthesia. Importantly, 600 μg misoprostol via the vaginal route in the midtrimester did not alter maternal haemodynamics as assessed by transthoracic bioimpedance measurements. A recent editorial points out that side effects, in particular shivering and hyperpyrexia, are dose dependent, and cautions that the administration of a therapeutic dose of misoprostol after an initial prophylactic dose may be harmful. A recent systematic review reports 11 maternal deaths during five trials. Eight of these women received 600 μg or more of misoprostol and 3 were controls. Further research will elucidate whether a dose of 400 μg is safer than 600 μg as prophylaxis, and whether doses as high as 800 μg are required for treatment of PPH.

Nausea and vomiting

Nausea and vomiting can make caesarean delivery under SA unpleasant. Contributing factors have been reviewed,[6] and uterotonic drugs are frequently causative. The incidence of nausea has been reported as 29%, and vomiting 9% after a bolus of 5 IU of oxytocin, and 10% after 250 μg of intramyometrial 15-methyl prostaglandin F2α during elective caesarean delivery under SA. A high incidence of 46% of nausea or vomiting has been reported after 0.5 mg ergometrine IV. A recent Cochrane review of the obstetric literature suggested that for the prevention of PPH > 1000 mL, syntometrine had a similar efficacy to oxytocin, but was associated with a 5-fold increase in nausea, vomiting and hypertension.

Summary: A plan of action
Uterotonic drugs remain an important intervention in the prevention of uterine atony. In addition, these agents are essential adjuncts to aggressive resuscitation and surgical management of PPH during CS.[7,8] Current evidence is that oxytocin remains the uterotonic of first choice. There are few definitive studies upon which to base a protocol for recommendations for dosing of oxytocin or second-line uterotonics. Recommendations vary considerably from country to country. The following is a reasonable protocol, based upon current literature:

A Prophylaxis of uterine atony

- In healthy parturients at low risk for uterine atony, a bolus dose of 3 IU IV over 30 seconds, repeated twice at three minute intervals, is reasonable. This may be followed by a continuous infusion containing 20 IU oxytocin at 100-150 mL/hour.
- In resource constrained environments where oxytocin infusion cannot be guaranteed, consider intramuscular syntometrine (0.5 mg ergometrine plus 5 IU oxytocin).
- In patients with preeclampsia, oxytocin infusion should be at the lowest required rate, to avoid fluid retention. Administration of oxytocin in advanced cardiac disease requires further investigation. Slow infusion of small doses would seem prudent, with special care in patients with pulmonary hypertension.
- Carbetocin is a useful long-acting oxytocin analogue, commonly administered in a dose of 100 μg IV or IM.

B Management of established uterine atony / postpartum haemorrhage

- In view of down-regulation of oxytocin receptors following prior exposure to oxytocin, there should be a low threshold for the use of uterotonics with a different mechanism of action, such as ergometrine, and prostaglandins F2α and E1, in cases of established uterine atony.
- Ergometrine 0.1-0.5 mg IV slowly or 0.5 mg IM, in the absence of contraindications, is the currently recommended second line therapy, but many practitioners use doses as low as 0.05 mg IV at CS.
- 15-methyl prostaglandin F2α, 250 μg – 500 μg is used intramyometrially as a last resort, but this agent is not licensed for administration via this route. Current guidelines stipulate that 250 μg should be given IM, repeated every 15 minutes to a maximum of 8 doses, as the last resort.
- In preeclampsia, the best second line agent is misoprostol, since ergot alkaloids are contraindicated, and prostaglandin F2α may be hazardous. The optimal dose and route of administration of misoprostol for the prophylaxis and management of PPH, remain to be established. Currently, expert opinion advises regimens such as 200 μg sublingually, together with 400 μg rectally.
- During established PPH, the infusion of oxytocin 40 IU/500 mL at 125 mL/hour has been recommended, accompanied by effective resuscitation and the co-administration of second line agents as described.
- Extreme care is required in the use of oxytocin in haemodynamically unstable patients.

B: Tocolytic drugs

Preterm delivery, defined as delivery before 37 weeks' gestation, occurs in 5-13% of pregnancies. Tocolysis per se does not prevent preterm delivery or improve neonatal outcomes, but allows for the initiation of other treatments or interventions that can improve outcomes. Short term tocolysis, typically <48 hours, in patients with a viable fetus of gestational age <34 weeks, allows for the administration of corticosteroids to for fetal lung maturity, and/or for urgent transfer to the referral unit or operating theatre. Tocolysis may also be given, for example, before external cephalic version, in the event of tachysystole with fetal heart changes, or to allow for magnesium sulphate administration in preterm infants, for neuroprotection. There are obstetric contraindications to tocolysis, including preeclampsia, and haemorrhage with haemodynamic instability.
Analyses comparing the use of available agents, calcium channel blockers, B agonists, cyclo-oxygenase inhibitors, atosiban, nitric oxide donors, progesterone, and magnesium sulphate with placebo as primary tocolysis, show that all of these agents have shown effect in prolonging pregnancy for up to 48 hours, but there is no beneficial effect on neonatal morbidity or mortality. Magnesium is not efficacious as a tocolytic agent, but may be administered for fetal neuroprotection before 32 weeks' gestation, or commonly for seizure prophylaxis in preeclampsia. Maintenance tocolysis is of no benefit, and may be harmful.[9,10]

Calcium channel blockers (CCB)

Lowering of intracellular calcium reduces MLCK activity. In many units, CCB, typically nifedipine 20 mg orally 6 hourly, are the most commonly used agents for preterm labour, because of the low incidence of side effects, which include hypotension, nausea, flushing, and occasionally pulmonary oedema.

β Agonists

Increased β 2 receptor activity increases intracellular cAMP and lowers intracellular calcium levels, reducing MLCK activity. These agents, ritodrine, terbutaline and salbutamol, have become less popular, due to maternal side effects, which include hypotension, tachycardia, arrhythmias, hypokalaemia and hyperglycaemia; however the overall incidence is low. In our unit, salbutamol 250 µg is typically administered IV as part of intrauterine fetal resuscitation, to inhibit labour contractions in a patient in labour and at risk of uterine rupture, for external cephalic version, or if there is cord prolapse in the presence of active labour.

Cyclo-oxygenase inhibitors

These agents, of which indomethacin is most commonly used, are inhibitors of prostaglandin synthesis. This is the only class of tocolytics which decreases the preterm birth rate at <37 weeks' gestation. During such short term use, only minor complications such as nausea and heartburn are noted. Transient ductal constriction has been noted in 50% of fetuses between 26- and 31 weeks' gestation. Used in short courses in very pre-term infants, closure of the ductus arteriosus is relatively uncommon (5-10% until 32 weeks, rising to 50% thereafter). It is usually reversible with early recognition and discontinuation of the agent. Oligohydramnios may also occur due to decreased fetal urine output, possible due to an enhanced anti-diuretic hormone effect. There may be an increased risk of necrotising enterocolitis and intraventricular haemorrhage.

Oxytocin receptor antagonist

The oxytocin receptor antagonist atosiban is not yet FDA approved in the US, but widely used in many countries. There are no apparent advantages over the established agents.

Nitric oxide donors

Once again, no advantages have been shown for the use of transdermal nitroglycerin.

Progesterone

There is insufficient evidence for the benefits of the use of progesterone in acute tocolysis, but some evidence that progesterone may have a role in maintenance tocolysis.

Conclusion

Overall, the anaesthesiologist should have a good understanding of the pharmacology and therapeutic range of drugs administered to alter uterine tone, since uterotonic agents have a crucial role to play in obstetric haemorrhage, and the use of tocolytic drugs impacts significantly on anaesthesia practice.

References

1. Dyer RA, van Dyk H, Dresner A. *The use of uterotonic drugs during caesarean section.* Int J Obstet Anesth 2010; 19:313-319.
2. Arias F. *Pharmacology of oxytocin and prostaglandins.* Clin Obstet Gynecol 2000; 43:455-68.
3. Tsen LC, Balki M. *Oxytocin protocols during caesarean delivery: Time to acknowledge the risk/benefit ratio?* Int J Obstet Anesth 2010; 19:243-245.
4. Farina Z, Fawcus S. *Oxytocin-ensuring appropriate use and balancing efficacy with safety.* S Afr Med J 2015; 105:271-274.
5. Meshykhi, Nel MR, Lucas DN. *The role of carbetocin in the prevention and treatment of postpartum haemorrhage.* Int J Obstet Anesth 2016; 28:61-69.
6. Balki M, Carvalho JC. *Intraoperative nausea and vomiting during cesarean section under regional anesthesia.* Int J Obstet Anesth 2005; 14:230-41.
7. http://www.rcog.org.uk *Prevention and treatment of postpartum haemorrhage*. Royal College of Obstetricians and Gynaecologists' Green-top Guideline no.52, 2009.
8. WHO Guidelines for the prevention and treatment of PPH (2012), at http://apps.who.int/iris/bitstream/10665/75411/1/9789241548502_eng.pdf?ua1/41.
9. Bolden JR. *Acute and chronic tocolysis.* Clinical Obstetrics and Gynecology 2014; 57:568-578.
10. Navathe R, Berghella V. *Tocolysis for acute preterm labor: Where have we been, where are we now, and where are we going?* Am J Peritol 2016; 33:229-235.

Transitional Foetal Physiology Adaptation

Dr Heidi Meyer

Specialist
Division of Paediatric Anaesthesia
UCT Dept of Anaesthesia & Perioperative Medicine

RESPIRATORY

Lung aeration

Lung aeration occurs rapidly, at a rate of approximately 3 mL/kg per second during inspiration. Two primary mechanisms aid clearance of amniotic fluid from the lungs (*Hooper, 2010)*;

1. **Transpulmonary pressure gradient**: The first inspiratory effort produces subatmospheric pressures (~50 cmH2) in the intrapleural and interstitial spaces. This drives the liquid into the interstitial tissue compartment from which it is gradually cleared by the pulmonary circulation and lymphatics. This has been shown to account for **>95% of lung clearance** of fluid in rabbit models (*te Pas, 2008*).

2. **Mechanical forces** play a minimal part in lung clearance during labour, but liquid still fills the airways until the newborn takes its first breath.

Establishment of FRC

The normal FRC (30 ml/kg) body weight is usually achieved within hours of birth, taking 2-3 hours in vaginally delivered term infants and 5-6 hours in infants delivered by caesarean section (*te Pas, 2008*).

Air-trapping

The first breaths tend to be deeper and longer than subsequent breaths and are characterized by a short deep inspiration followed by a prolonged expiratory phase. Essentially an FRC is established by inhaling larger volumes than they exhale, resulting in the accumulation of FRC with each breath.

Expiratory braking manoeuvres

Expiratory flow is interrupted by a period of low or zero flow, resulting in a short expiratory flow peak or multiple expiratory flow peaks. This expiratory breaking is achieved by **post-inspiratory activity of the diaphragm** counteracting the passive recoil of the lung and by **adduction of the glottis** during expiration.

Surfactant

Surfactant produced by type II alveolar cells plays an important role in reduction of surface tension resulting in increased lung compliance and enabling more uniform aeration of the alveoli.

Amiloride-sensitive epithelial sodium channels

Adrenaline and ADH are released by the foetus during labor. These hormones activate amiloride-sensitive epithelial Na channels (ENaCs), which switch from facilitated Cl^- secretion to active Na^+ reabsorption from the lung lumen into the interstitial tissue compartment leading to liquid reabsorption. There appears to remain some debate as to the significance of role of ENaCs in lung aeration and maintenance of FRC.

Initiation of ventilation and apnoea

Irregular breathing activity already occurs in the foetus during REM sleep. During normal transition the initiation and maintenance of breathing occurs due to the combined stimuli of cord clamping (and

the probable removal of rapidly catabolized prostaglandins that suppress breathing), tactile and cold stimuli that act centrally, and changes in PCO_2 and PO_2 levels in the blood.

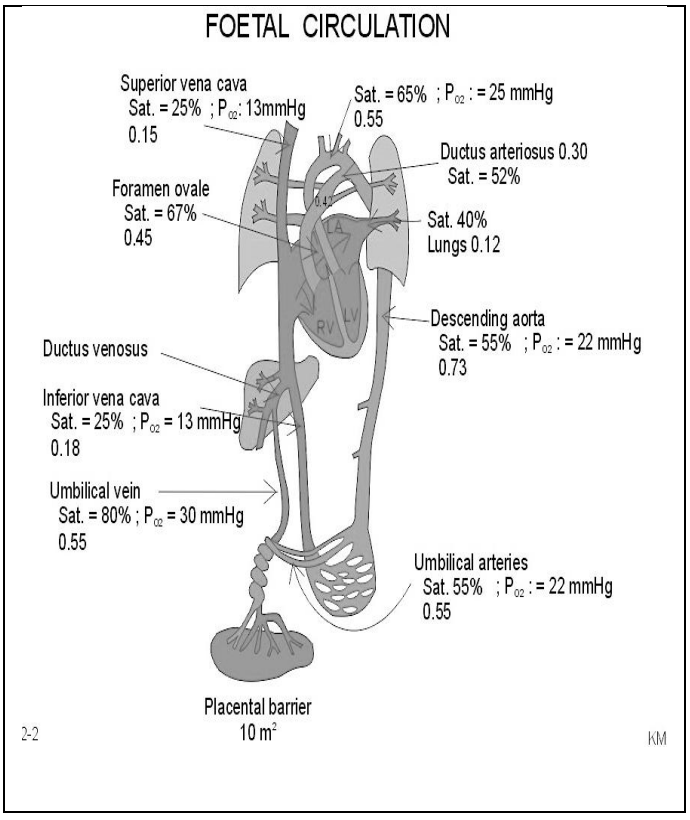
Asphyxia in the neonate will initially result in an increase in respiratory rate and then apnoea. This primary apnoea will respond to stimulation. If asphyxia continues after the episode of primary apnoea, the neonate will respond with a period of gasping respirations and then enter a pre-terminal secondary apnoea, which will require active resuscitation.

CARDIOVASCULAR

Foetal circulation

Oxygenated blood from the placental bed returns to the foetus via the umbilical vein. Approximately 50% of the oxygenated blood returning from the placental bed is then shunted through the ductus venosus, bypassing the liver, and enters the right atrium. This relatively well oxygenated blood is streamed by the Eustachian valve through the right atrium and across the foramen ovale into the left atrium. The blood from the left atrium passed through the mitral valve, into the left ventricle and is ejected into the root of the aorta to then preferentially supply the brain and upper limbs with the relatively well oxygenated blood. The relatively deoxygenated blood returning from the SVC combined with blood from the coronary sinus passes through the right atrium into the right ventricle. The right ventricle ejects blood into the pulmonary artery, approximately 11% of flow then passes through the high resistance pulmonary vascular circulation, whilst the remainder of the blood passes through the ductus arteriosus and into the descending aorta where it supplies the lower half of the body and supplies the umbilical arteries to be re-oxygenated at the placental bed.

The flow from the two unequal parallel circulations is described as the percentage of combined cardiac output (CCO), which is the combined output of the left and right ventricles. The human fetal combined cardiac output at term has been estimated to be approximately 500 mL per minute per kilogram. The distribution of flow within the foetus expressed as a percentage of CCO is detailed in the table below.



Key points

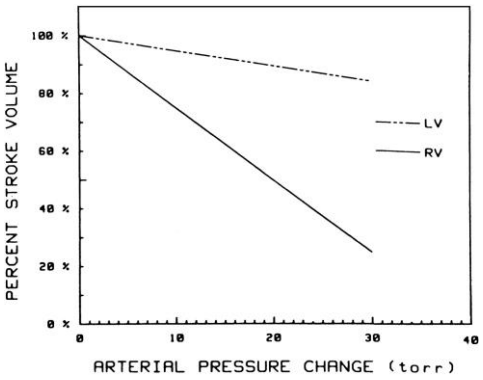
- Two parallel circulations
- Gas exchange occurs via a low resistance placental circulation
- Dependent on intra-cardiac and extra-cardiac shunts
- High resistance pulmonary circulation

Distribution of flow in foetus expressed as percent of combined cardiac output	
Right cardiac output, %	59
Left cardiac output, %	41
Ductus arteriosus blood flow, %	46
Pulmonary blood flow, %	11
Foramen ovale blood flow, %	33

(*Mielke G, 2001*)

Right & left ventricles

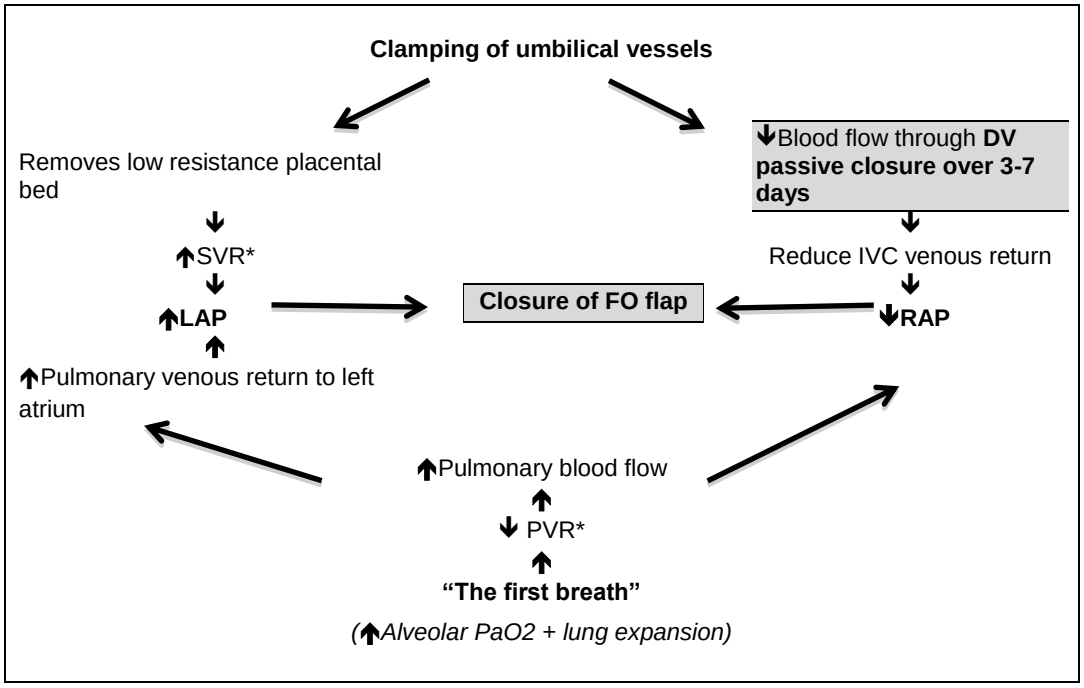
In the embryo and the foetus the right ventricle is the dominant ventricle. The right ventricle has a larger chamber volume, a larger radius-to-wall thickness ratio, a higher free wall stress, and greater sensitivity to increases in arterial pressure. The right ventricle performs more work and has higher metabolic requirements in the face of stresses when arterial pressure is increased. The stroke volume of the right ventricle decreases significantly in the face of increased afterload, compared with the left ventricle (*Greyson, 2010*).



Cardiovascular adaptation and changes at birth

At birth, the following must occur:

- Significant drop in PVR and increase in pulmonary blood flow
- Closure of 3 shunts; **ductus venosus (DV), foramen ovale (FO) and ductus arteriosus** (DA)
  Note: this will only be functional closure initially
- Transition from parallel circulation to **circulation in series** = adult circulation



- *↑SVR + ↓ PVR ➔ Reversal of flow through DA ➔ L ➔ R shunt
- ↑ PaO_2+ ↓ Prostaglandin E2 ➔ **Functional closure of DA (24-48 hrs)**
- Permanent structural closure after 4-8 weeks via endothelial destruction and sub-intimal proliferation

Changes in pulmonary vascular resistance

At birth there is a rapid decline in pulmonary vascular resistance secondary to lung expansion and an increase in the PO_2. This results in a substantial increase in pulmonary blood flow. The right ventricular pressures still exceed systemic pressures, but will begin to fall over the few hours to days following birth. Pulmonary pressures tend to fall below systemic pressures by approximately 3 weeks of age (*Greyson, 2010)*.

Right & left ventricle adaptation

After birth and closure of the foetal shunts the circulation is now in series and each ventricle will now receive 100% of the cardiac output. The closure of the ductus arteriosus means right ventricular output will now flow through the pulmonary artery into the low resistance pulmonary circulation. The right ventricle atrophies and the left ventricle will begin to hypertrophy as it assumes the role of the sub-systemic ventricle ejecting against the increased systemic vascular resistance. By 3 – 6 months the classical left ventricular dominance of adulthood is established.

Persistent foetal circulation

In the presence of certain stimuli, the pulmonary arterioles will constrict and lead to an increase in pulmonary vascular resistance (PVR). The increase in PVR results in reversal of shunting to a right to left shunt through the foramen ovale and ductus arteriosus, leading to hypoxia and a pathophysiological state termed persistent foetal circulation. These stimuli include; hypoxia, hypercarbia, acidosis, hypothermia, pain, and inadequate depth of anaesthesia.

HAEMATOLOGICAL

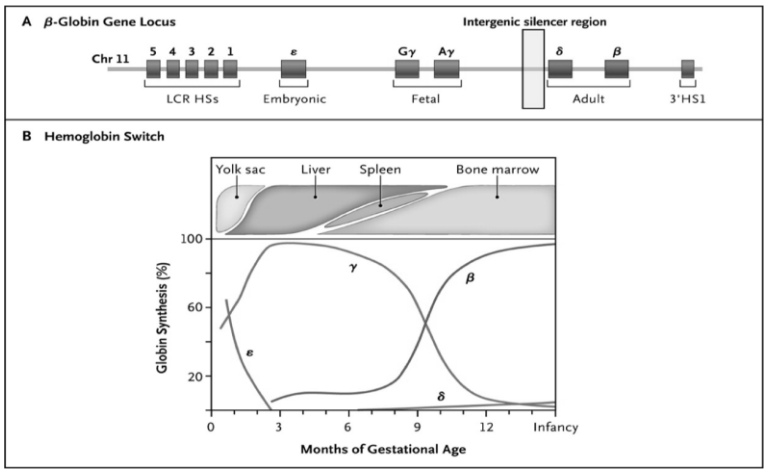
Erythropoiesis

Erythropoiesis occurs at different sites throughout development. Erythropoiesis is yolk-sac derived during early embryonic development, transitioning to the foetal liver mid-way through the first trimester. As birth approaches the bone marrow becomes the dominant site of erythropoiesis.

Globin synthesis

During foetal development **two switches** occur in the expression of genes from the beta-globin cluster.

1. Early gestation: switch from embryonic to foetal globins
2. Around the time of birth: switch from foetal to adult globins



A. Beta-globin gene locus

B. Foetal haemoglobin (HbF, α2γ2) constitutes the major hemoglobin during fetal life. It is gradually replaced by adult hemoglobin (HbA, α2β2) during infancy and HbF levels rapidly decline by approximately 6-9 months of age.

Adults retain low-level expression of HbF (0.1% and 1% total hemoglobin), with a small proportion of erythrocytes containing measurable HbF.

Figure x. Haemoglobin synthesis from the embryo to adult

Foetal haemoglobin structure

In adult haemoglobin there are six positively charged amino acids within the pocket that tightly bind the negatively charged 2,3-BPG molecule in the tense (T) state. Thus stabilising the deoxygenated haemoglobin molecule and reducing the affinity for oxygen, reflected as a shift of the oxygen dissociation curve to the right. Foetal haemoglobin differs from adult haemoglobin due the presence of **gamma subunits versus beta subunits**. This results in an alteration of charge in the 2,3-BPG binding pocket at the center of the haemoglobin molecule.

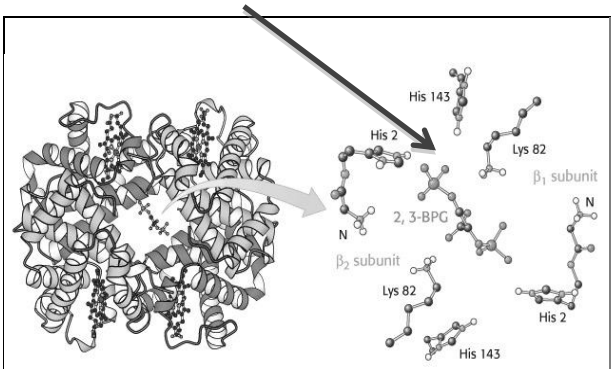
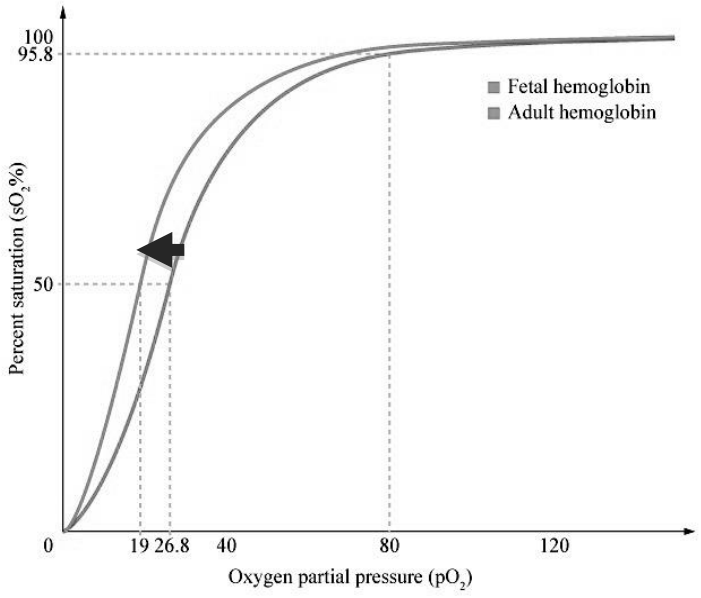


Figure. Adult haemoglobin molecule and amino acid configuration of the 2,3 – DPG binding pocket

In HbF, the presence of the gamma globulin chains, results in the replacement of two positively charged histidine amino acids with two neutrally charged serine amino acids. The overall reduction in the positive charge within the pocket results in **reduced affinity for the 2,3-BPG**. Therefore the HbF has an increased affinity for oxygen compared with HbA.

Foetal haemoglobin oxygen dissociation curve



The '*P*50' is the partial pressure of oxygen in blood in which haemoglobin is 50% saturated.

Change in P50 value therefore describes whether the curve has shifted **to the left** (increased affinity for O_2) or to the right (reduced affinity for O_2).

The decreased affinity for 2,3-BPG resulting in HbF having an increased affinity for oxygen compared with HbA, resulting in a left shift of the oxygen dissociation curve.

***P*50 for HbA is 3.5 kPa (26.8 mmHg)**

***P*50 for HbF is 2.5 kPa (19 mmHg)**

Physiological anaemia

After birth the increased PaO_2 results in down regulation of erythropoietin and subsequent reduction in red blood cell mass. The red blood cell mass will decrease until it reaches a critical level where oxygen delivery is inadequate for demand. This is referred to as the nadir and usually occurs at approximately 8-12 weeks post delivery and reaches a level of approximately 9.0 – 11.0 $g.dl^{-1}$. This nadir occurs earlier in preterm infants, at 6-9 weeks, and results in lower haemoglobin levels of 7.0-9.0 $g.dl^{-1}$.

Coagulation

Coagulation factors do not cross the placenta; however, factors V, VIII and XIII and fibrinogen are at adult concentrations before birth.

The vitamin K-dependent clotting factors (procoagulant; II, VII, IX, X and anticoagulant; Protein C & S) are reduced compared with normal adult values because of a lack of vitamin K stores and immature hepatocyte function. The haemostatic system is not fully mature until 3 to 6 months of age (*Pichler E & Pichler L, 2008*).

Normal haematological ranges for term and pre-term babies from the UK Blood Transfusion and Tissue Transplantation Services Handbook

	Term	Preterm	Adult
Haemoglobin g/l	140−240	140−240	115−180
Platelets x 10^9/l	150−450	150−450	150−400
PT (sec)	10−16	11−22	11−14
APTT (sec)	31−55	28−101	27−40
TT (sec)	19−28	19−30	12−14
Fibrinogen g/l	1.7−4.0	1.5−3.7	1.5−4.0

Pre-transfusion testing for blood transfusion up to age 4 months

Whilst the ABO antigen is fully developed at birth, the newborn baby does not produce ABO antibodies up to 3-6 months of age. The ABO antibodies present in the serum of newborn babies are derived from mother's blood due to placental transfer. This has a clinical implication on pre-transfusion testing for blood transfusion. In the neonate the required testing includes ABO, Rh(D) typing and an antibody screen. The antibody screen is performed to detect red cell antibodies and can be performed using either a neonatal or a maternal blood specimen. If the initial antibody screen is negative, a full cross match is not required and packed red blood cells can be requested without repeating the group & screen up to four months of age.

Oxygen transfer to the foetus

Factors affecting oxygen transfer from mother to foetus

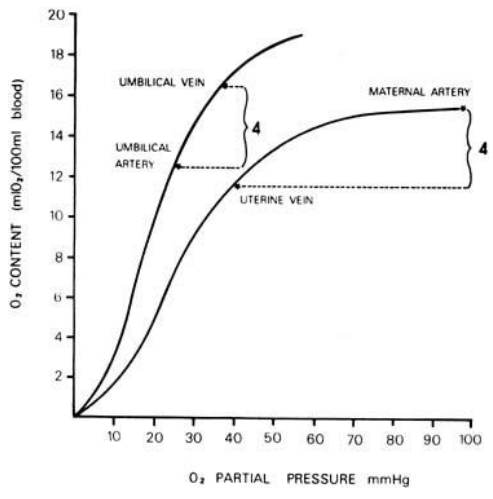
- Intervillus blood flow
- Foetal placental blood flow (~55% CVO)
- Oxygen tension in maternal blood
- Oxygen tension in foetal blood
- Oxygen affinity of maternal haemoglobin
- Oxygen affinity of foetal haemoglobin
- Haemoglobin concentration of maternal blood
- Haemoglobin concentration of foetal blood
- Double Bohr effect
- Placental oxygen consumption

Maternal vs foetal oxygen content

The differences between maternal and fetal O2 content, as represented in the PO_2 curves shown, are due to differences between maternal and fetal hemoglobin:

1. Fetal hemoglobin's greater affinity for oxygen shifts the curve to the left

2. Greater fetal hemoglobin concentration makes the curve steeper



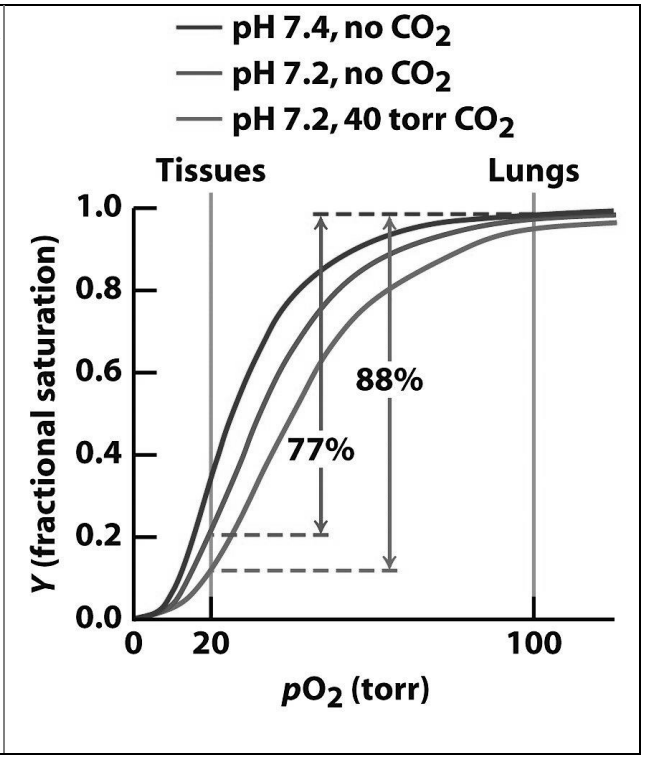
Double Bohr effect

Bohr effect

- Haemoglobins oxygen binding capacity is inversely proportional to the acidity and carbon dioxide concentration
- Promotes oxygen release at the tissue level
- Promotes oxygen uptake in the lungs

Double Bohr effect

- Helps to increase foetal oxygenation
- CO_2 diffuses across placenta from foetal to maternal circulation
- CO_2 decreases in the foetal circulation, shifting the curve to the left
- CO_2 increases in the maternal circulation, shifting the curve to the right



CONCLUSION

The adaptation for extra-uterine life at birth results in dramatic and rapid changes from the foetal to the neonatal physiology.

These include rapid changes to the respiratory and cardiovascular systems as well as more gradual changes to the haematological, renal, hepatic and nervous systems. Changes also occur in thermoregulation, body fluid composition and nutrition.

References

*Underlined articles are open access

1. **Hooper ST**, te Pas AB, Lewis RA, Morley CJ . **Establishing FRC at birth**. *Neoreviews*. 2010;11(9):474-483
2. **te Pas AB**, Davis PG, Hooper SB, Morley CJ. **From liquid to air: breathing after birth**. *Journal of Pediatrics 2008;152(5):607-11*
3. **Greyson CR**. **The right ventricle and pulmonary circulation: Basic concepts**. *Rev Esp Cardiol*. 2010;63(1):81-95
4. **Mielke G** & Benda M. **Cardiac Output and Central Distribution of Blood Flow in the Human Fetus**. *Circulation*. 2001;103:1662-1668
5. **Pichler & Pichler L. The neonatal coagulation system and the vitamin K deficiency bleeding** – a mini review. *Wien Med Wochenschr*. 2008;158(13–14):385–39

Recommended reading: **Sharma AS**, Ford S, Calvert J. **Adaption for life: a review of neonatal physiology**. *Anaesthesia and Intensive Care Medicine*. 2010;9(3):85-90

Iron- Physiology & Pharmacology

Dr Dominique van Dyk

UCT Department of Anaesthesia & Perioperative Medicine

PHYSIOLOGY OF IRON

The importance of iron

Iron is the fourth most abundant element of the earth's crust and the second most abundant metal. It is also an essential nutrient required by every human cell.

Its atomic structure gives rise to a number of biochemically useful properties, including the unusual capacity to both donate and accept electrons, and to reversibly bind to ligands such as oxygen and nitrogen. It can exist in various oxidation states (from **-2** to **+6**), the principal being **Fe2+** and **Fe3+.** The body exploits the unique properties of iron by incorporating it into hundreds of different enzymatic and non-enzymatic proteins that are crucial to a wide range of physiological functions:

FUNCTION	PROTEIN
Oxygen transport and storage	**Haemoglobin** in red blood cells transports oxygen in the blood, and **myoglobin** stores oxygen in muscles
Oxygen homeostasis	An iron-dependent **prolyl hydroxylase** plays a critical role in the physiologic response to hypoxia
Electron transport and energy production	**Cytochromes** and **dehydrogenases** are crucial components of mitochondrial electron transport for ATP synthesis
Metabolism and detoxification	**Cytochromes** are also involved in the metabolism of biological molecules, drugs and pollutants
Antioxidant activity	**Catalase** and **peroxidases** metabolise hydrogen peroxide to reduce the risk of oxidative cellular damage
Beneficial pro-oxidant activity	**Myeloperoxidase** synthesizes reactive oxygen species with neutrophils to aid bacterial cell killing
DNA synthesis	**Ribonucleotide reductase** converts ribonucleotides to deoxyribonucleotides.
Neurological function	Cofactor in the synthesis of **neurotransmitters** and **myelin**

The biochemical reactivity of iron – namely its ability to both donate and accept electrons – means that it can be harmful when present in high concentrations. Unbound iron can catalyse the formation of reactive oxygen species that cause intra- and extra-cellular damage. Thus, when not bound to functional proteins (functional iron), free iron is sequestered by the iron transport protein transferrin (transport iron) or stored by ferritin or as haemosiderin (storage iron). Iron levels must be tightly regulated to prevent damage caused by either too much or too little iron.

Distribution of iron

Body iron content is approximately 4-5 g in well-nourished populations, which corresponds to a concentration of about 50 mg iron/kg body weight in men and about 40 mg iron/kg BWT in women. Approximately 60% is present in the form of haemoglobin in red blood cells
Most of the rest is contained in ferritin complexes that are found in all cells, but most commonly in the bone marrow, liver, and spleen. The ferritin molecule is composed of an apoferritin protein shell, composed of 24 polypeptide subunits (of 2 types, named L & H) with an overall molecular weight of approximately 500 kDa, enclosing a crystalline ferrihydrite core of up to 4,500 ferric iron atoms. The liver's stores of ferritin are the primary physiologic source of reserve iron in the body. Haemosiderin is an aggregated, iron-rich ferritin degradation product with very low apoferritin content, found in some phagocytic cells (such as macrophages). Its iron content is not readily available.

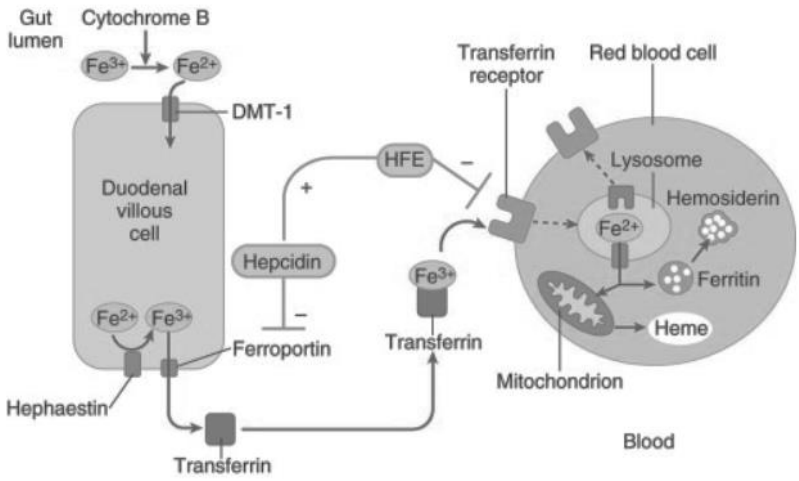
TYPE OF IRON	CONCENTRATION (MG IRON/KG BODY WEIGHT)	
	Men (% of total)	Women (% of total)
Functional Iron		
Haemoglobin	31 [62]	28 [70]
Myoglobin	5 [10]	4 [10]
Haem enzymes	1 [2]	1 [2.5]
Non-haem enzymes	1 [2]	1 [2.5]
Transport iron		
Transferrin	<1 (0.2 [0.4])	<1 (0.2 [0.5])
Storage iron		
Ferritin	8 [16]	4 [10]
Haemosiderin	4 [8]	2 [5]
Total	~50	~40

Systemic and cellular regulation of iron

Human iron homeostasis is regulated at two different levels, the **systemic** level and the **cellular** level, by two very distinct control systems.
Systemic iron levels are controlled by the hepcidin/ferroportin system- the regulated absorption of dietary iron by enterocytes, the cells that line the interior of the intestines, is balanced against the uncontrolled loss of iron from epithelial sloughing, sweat, injuries and blood loss. In addition, systemic iron is continuously recycled.
Control of **cellular iron levels** is accomplished using mechanisms involving the expression of particular iron regulatory and transport proteins, described as the IRE/IRP system.
It seems very likely that there is higher level coordination between the systemic and cellular systems, and future work will define the detail.

Plasma iron
The key to systemic iron supply and homeostasis lies in the regulation of adequate plasma iron levels. Iron circulates in plasma bound to the 80-kDa glycoprotein **transferrin.** The liver is the main site of transferrin synthesis but other tissues and organs, including the brain, also produce transferrin. The main role of transferrin is to deliver iron from macrophages and duodenal absorption to all tissues. Transferrin has two high-affinity binding sites for Fe3+.
Transferrin binding (1) maintains iron in a soluble form, (2) serves as a major vehicle for iron delivery into cells (via the transferrin receptor, TfR1), and (3) limits the generation of toxic radicals.
Plasma transferrin is normally about 30% saturated with iron. A transferrin saturation <16% indicates iron deficiency, whereas >45% saturation is a sign of iron overload. When the saturation exceeds 60%, *non-transferrin-bound iron* begins to accumulate in the circulation and to damage parenchymal cells.
The homeostatic system thus has to maintain transferrin saturation at physiological levels, responding to signals from pathways that consume iron (such as erythropoiesis) and sending signals to the cells that supply iron to the blood stream.
Iron is released into the circulation from:
(1) *macrophages of the reticuloendothelial system*, which internally recycle 20–25 mg of iron from senescent erythrocytes per day, and
(2) *duodenal enterocytes*, which absorb 1–2 mg of dietary iron per day.
Hepatocytes play a dual role in systemic iron metabolism: they are the major site of iron storage and they secrete the regulatory hormone **hepcidin**. Hepcidin orchestrates systemic iron fluxes and controls plasma iron levels by binding to the iron exporter **ferroportin** on the surface of iron-releasing cells, triggering its degradation and hence reducing iron transfer to transferrin. Inherited and acquired disorders that alter hepcidin production therefore cause iron deficiency (high hepcidin levels) or iron overload (hepcidin deficiency).
Assessing the concentration of serum ferritin is a clinically useful measure of iron storage. Iron is stored inside cells as ferritin. Small amounts of ferritin are secreted into the plasma, proportional to the

amount within the cells. Low serum ferritin levels indicate depleted stores, whereas increased levels may indicate iron overload. Inflammatory conditions (or infections, cancer, and liver disorders) can also increase serum ferritin. The physiological function(s) of serum ferritin and its source (that is, whether it is derived from damaged cells or actively secreted by a regulated mechanism) still remain to be defined. Serum ferritin is predominantly composed of L chain subunits, partially glycosylated, and iron-poor.



Iron Transport, from *Hematology- A pathophysiologic approach*, Hudnall ed.

Iron Absorption

Most of the iron in the diet is in the ferric (Fe3+) form, whereas it is the ferrous (Fe2+) form that is absorbed. Gastric secretions dissolve the iron and permit it to form soluble complexes with ascorbic acid and other substances that aid its reduction to the Fe2+ form. The importance of this function in humans is indicated by the fact that iron deficiency anaemia often complicates partial gastrectomy. Inorganic dietary iron is absorbed at the brush border of duodenal enterocytes via the **divalent metal transporter 1 (DMT1)**. A membrane-associated ferrireductase**, duodenal cytochrome B**, must first reduce any iron in its oxidized state before this absorption can occur. Haem iron is absorbed by a **haem importer protein** of uncertain identity, and is released intracellularly from its porphyrin by haemoxygenase, mainly by the inducible haemoxygenase 1 (HOX1). Within the enterocyte, some of the Fe2+ is converted to Fe3+ and bound to ferritin. The rest of the cytosolic iron binds to the basolateral Fe2+ transporter, **ferroportin**, and is exported to the interstitial fluid. Enterocytic iron export through ferroportin requires **hephaestin**, a membrane-bound multicopper oxidase homologous to serum caeruloplasmin, which oxidases Fe2+ to Fe3+ for loading onto transferrin. In the presence of hepcidin, ferroportin is internalized and degraded. Thus, iron exportation is blocked. Inversely, in the absence of hepcidin, ferroportin is maintained on the cell membrane, and iron transportation is facilitated.

Because iron cannot be excreted from the organism in a regulated way, iron absorption represents the critically controlled process. Normally, only 1–2 mg of iron per day are absorbed to compensate for iron losses, for example by sloughing of intestinal epithelial cells, desquamation of skin and urinary cells, blood loss, or sweat. Iron absorption can be enhanced when the needs are higher (for example, because of increased erythropoiesis or pregnancy) and suppressed in iron overload. The lack of an active mechanism for iron excretion explains the development of iron overload when the regulation of iron absorption is defective or bypassed (e.g. blood transfusions & IV iron therapy).

Iron Recycling

Less than 10% of the daily iron needs are met by intestinal absorption, and the rest is covered by macrophages that recycle iron internally to maintain plasma iron levels. The bulk of recycled iron (about 25 mg/day) is used by erythroblasts for haemoglobin synthesis. The amount of plasma iron is just over 10% of the amount used daily, which means that plasma iron is turned over many times each day.

Macrophages phagocytose aged or damaged erythrocytes and catabolize haem using haemoxygenase. NRAMP1 (natural resistance-associated macrophages protein 1), a divalent metal transporter homologous to DMT1, is expressed within phagolysosomal membranes and participates in iron export from phagocytic vesicles. Export of ferrous iron from macrophages occurs via ferroportin.

Because it has such a central role in systemic iron homeostasis, ferroportin expression in macrophages is closely regulated:
1) ferroportin transcription is induced by erythrophagocytosis and haem iron,
2) translation of ferritin is regulated by the IRE/IRP system
3) protein stability of ferritin is regulated by hepcidin
Ferroportin-mediated iron export is coupled to the function of the multicopper oxidase caeruloplasmin, a protein synthesized and secreted by the liver. Caeruloplasmin-deficient humans show hepatocyte and macrophage iron accumulation.

Systemic Iron Homeostasis: More about the Iron Hormone Hepcidin
Control of iron balance in the whole body requires communication between sites of uptake, utilization and storage. Hepcidin (encoded by the HAMP gene) is the central regulatory molecule of systemic iron homeostasis. It is a defensin family member with strong links to innate immunity. The bioactive, mature 25 amino acid peptide is generated from an 84 amino acid prepropeptide by furin cleavage in the liver. Hepcidin is secreted from hepatocytes and circulates in plasma bound to α2-macroglobulin. Hepcidin clearance occurs via the kidney or by codegradation with ferroportin.
Hepcidin binds to ferroportin, triggers its internalization, ubiquitination, and subsequent lysosomal degradation.). In the duodenal enterocyte, hepcidin-dependent regulation of ferroportin reduces dietary iron absorption, i.e. it is a negative regulator of iron absorption. In the macrophage (and possibly the hepatocyte), hepcidin activity attenuates cellular iron release by reducing ferroportin.
Hepcidin expression in hepatocytes is regulated by multiple, in part opposing signals, which include:
1) systemic iron availability (i.e. levels of diferric transferrin, Tf-Fe2): ↑Tf-Fe2 → ↑ hepcidin
2) hepatic iron stores: if high → ↑ hepcidin
3) erythropoietic activity: anaemia → ↓ hepcidin
4) hypoxia → ↓ hepcidin
5) inflammatory/infectious states: IL1 & IL6 → ↑ hepcidin
6) non-genetic iron overload → ↑ hepcidin
These different regulatory inputs are integrated transcriptionally.

Cellular Iron Homeostasis: The IRE/IRP System
Coordination of iron uptake, utilization, and storage to assure the availability of appropriate supplies and to prevent toxicity is as crucial on the cellular level as it is on the systemic level. In contrast to systemic iron metabolism, though, cellular iron traffic also involves regulated iron **excretion**.

Cellular Iron Uptake
Cells take up iron-loaded (diferric) transferrin, **Tf-Fe2,** via the high-affinity transferrin receptor, **TfR1**, as their major source of iron. The Tf-Fe2/TfR1 complex is internalized by clathrin-dependent endocytosis. Acidification in the endosomes, likely through the pump action of a Na^+-H^+-ATPase, triggers release of iron by conformational changes in both transferrin and its receptor. Freed iron is reduced from Fe3+to Fe2+ by metalloreductases for transport into the cytosol via the endosomal protein DMT1 (also found on the apical membrane of enterocytes where it mediates systemic iron absorption). The transferrin cycle is completed when the endosome returns to and fuses with the plasma membrane, returning apotransferrin to the circulation and TfR1 to the plasma membrane and allowing both molecules to start the cycle again.
There also seem to be transferrin-independent routes of iron uptake, and receptor-mediated endocytosis of other forms of protein-bound iron represents an additional means for specific cell types to take up iron. Serum ferritin can enter cells via at least 2 different ferritin receptors. Finally, specialized cells, e.g. enterocytes, are able to acquire iron in the form of haem. Cells also acquire haem indirectly. Macrophages obtain haem by phagocytosis and processing of senescent red blood cells. In plasma, haemoglobin and free haem arising from intravascular haemolysis are cleared by specific scavenger systems: haemoglobin forms a complex with haptoglobin that is delivered to reticuloendothelial cells via CD163-mediated endocytosis. Free plasma haem binds to hemopexin and the complex is endocytosed via the CD91 receptor present on the surface of macrophages, hepatocytes, and other cell types.

Cellular Iron Export
Iron export occurs from many cells but it is particularly important in cells that maintain plasma iron levels, i.e. macrophages and duodenal enterocytes. As mentioned before, ferroportin transports Fe2+ and acts in concert with either of the ferroxidases hephaestin (enterocytes) or caeruloplasmin (other cell types) that facilitate iron extraction from the ferroportin channel and subsequent loading onto plasma transferrin. Caeruloplasmin and hephaestin are both copper dependent & this explains the importance of the copper status for iron metabolism.

Erythroblasts: In order to avoid dangerous iron and/or haem excess (excess haem triggers apoptosis in proerythroblasts), erythroblasts store excess iron as ferritin or export it via **ferroportin**. Erythroblasts express a *ferroportin* messenger RNA (mRNA) isoform (1b) that lacks the 5′ IRE and thus evades potential translational repression by IRPs. This isoform is susceptible to hepcidin degradation and may provide erythroid precursors with a mechanism to respond to systemic iron availability. Additionally, erythroblasts have the capacity to export excess haem (for example, when globin synthesis is limiting).

Cellular Iron Storage
Iron from the cytoplasmic "labile iron pool" (LIP) that is not used by the cell or exported is stored within the nanocavity of ferritin heteropolymers made of 24 subunits of heavy and light chains. Ferritin provides cells with a means to lock up excess iron in a redox inactive form to prevent iron-mediated cell and tissue damage. It also serves as an iron store which can be mobilized by both proteasomal and lysosomal ferritin degradation. A few cell types (heart, testis, pancreas, kidneys) have homopolymeric (FtMt subunit) H-type ferritin present in their mitochondria which protects these organelles against iron-mediated toxicity. Unlike cytosolic ferritin, mitochondrial ferritin subunit expression is not (directly) controlled by the IRPs (see below).

Regulation of Cellular Iron Metabolism
While key aspects of systemic iron metabolism are regulated transcriptionally (hepcidin expression) and posttranslationally (ferroportin function by hepcidin), cellular iron homeostasis is regulated posttranscriptionally by iron regulatory protein 1 (IRP1) and IRP2. These two RNA-binding proteins interact with hairpin structures known as iron-regulating elements (IREs), which are present in the untranslated regions of target mRNAs for proteins involved in iron storage, utilization and export. IRP-binding to IREs responds to cellular iron levels.

Intracellular Iron Trafficking and Utilisation
How iron moves within cells remains poorly understood. In the cytoplasm, Fe2+ iron is directly bound to proteins such as ribonucleotide reductase, but most iron is transferred to mitochondria, where it is incorporated into **bioactive haem** and **Fe/S cluster prosthetic groups**. Iron is imported into mitochondria by the inner membrane protein **mitoferrin 1**; this process is facilitated by the ABCB10 protein, which is thought to stabilize mitoferrin 1. Potentially, iron may also be directly transported from endosomes into mitochondria by a "kiss-and-run mechanism" through a direct contact between both organelles, effectively bypassing the cytosol.
In the mitochondria, iron is delivered to the Fe/S cluster biosynthetic machinery, or inserted into protoporphyrin IX by ferrochelatase to form haem. To coordinate the synthesis of the haem precursor protoporphyrin IX with iron availability, δ-aminolaevulinic acid synthase 2, the erythroid-specific first enzyme of protoporphyrin IX synthesis, is post-transcriptionally regulated by iron via the iron-responsive element/iron-regulatory protein (IRE/IRP) system. Haem is exported from the mitochondria via a yet undefined mechanism for incorporation into proteins throughout the cell.

By making haem and Fe/S clusters, mitochondria represent the major subcellular site of iron utilisation and as such play a central role in the control of cellular iron metabolism. A current theory of how mitochondria influence cellular iron metabolism states that cells sense mitochondrial iron insufficiency via an Fe/S cluster-dependent factor and respond by increasing mitochondrial iron levels; a haem intermediate could also be involved. Diversion of iron to mitochondria depletes the cytosol, thereby stimulating IRP binding to IREs. This in turn increases cellular iron uptake (TfR1, DMT1) and diminishes iron storage (ferritin) and export (ferroportin), so that more iron becomes available. In erythroid cells, IRP activation furthermore inhibits haem synthesis to avoid the accumulation of toxic metabolic intermediates until mitochondrial iron sufficiency is restored.

IRON STUDIES

Measured indices include:

- ***Serum iron***- considerable daily variation; not very useful
- ***Serum transferrin (or total iron binding capacity, TIBC)*** - TIBC is a direct measure of the levels of transferrin; transferrin levels are reduced in inflammatory processes**.**

- ***Transferrin saturation*** - Level of transferrin saturation is particularly helpful in assessment of early stages of iron overload, with levels > 55% for males and > 50% for females indicative of iron overload (should be fasting level for more accurate assessment)
- ***Serum ferritin*** - Small amount of circulating serum ferritin reflects body iron stores; normal range 15 – 300 mcg/L (reference ranges vary depending on the method used). Levels < 15 mcg/L reflect absent / reduced iron stores. Elevated levels may reflect iron overload but will be increased in liver disease, inflammation or malignant disease. In the presence of inflammation, a level of > 100 mcg/l generally excludes iron deficiency
- ***Soluble transferrin receptor*** - Transferrin receptors are present on cell surfaces and are responsible for the internalization of transferrin resulting in intracellular release of iron. In the absence of adequate iron stores, expression of transferrin receptors increases. The amount of soluble transferrin receptor closely reflects iron stores and is not affected by inflammatory processes. Increased levels of soluble transferrin receptor are also found in conditions of increased red cell turnover (e.g. haemolysis)

Expected findings: "most helpful test" given in bold

	IRON-DEFICIENCY ANAEMIA *	**ANAEMIA OF CHRONIC DISEASE**	**IRON-DEFICIENCY & INFLAMMATION**	**ACUTE-PHASE RESPONSE**	**IRON OVERLOAD**
Serum iron	decreased	decreased	decreased	decreased	increased
Serum transferrin, TIBC	increased	decreased	decreased (low normal)	decreased	decreased or normal
Transferrin saturation	decreased	decreased	normal or decreased	decreased	**increased**
Serum ferritin	**decreased**	normal (>100 mcg/L)	"normal"	increased	**increased**
Soluble transferrin receptor	increased	**normal**	**increased**	**normal**	decreased

***Iron-deficiency anaemia:** Unless there is major bleeding, overt iron-deficiency anaemia develops progressively. The first sign of iron supply-demand mismatch is a reduction in serum ferritin, indicating that iron stores are being drawn on as iron is mobilised from the liver and reticuloendothelial system. However, serum iron, total iron-binding capacity (TIBC), and red cell morphology remain normal until iron stores are exhausted. Afterwards, serum iron levels decrease, while TIBC increases in an attempt to raise iron absorption. Dysfunctional erythropoiesis and microcytosis only occur once the transferrin saturation drops below 15%. Only then do anaemia and low haemoglobin levels develop.

PHARMACOLOGY OF IRON THERAPY

Oral iron supplements:

Indications: treatment of iron-deficiency anaemia; prevention of iron-deficiency anaemia in pregnancy; supplementation in regular blood donors to prevent subclinical iron deficiency.

Of course, the primary cause of the iron deficiency must be actively sought and corrected. The goal of iron supplementation is two-fold: to reverse anaemia and to replenish iron stores. The expected response to a course of oral iron is a reticulocytosis in 3-5 days, peaking after one week, followed within three weeks by a rise in haemoglobin. An increase in haemoglobin by 1g/dL after one month qualifies as an adequate response. The British Society of Gastroenterology suggests that oral iron therapy should be continued for three months after normalization of the haemoglobin levels to ensure that the iron stores are replenished.

Preparations:

The ferrous salts (including ferrous sulphate, ferrous gluconate, and ferrous fumarate) are the most commonly prescribed preparations. Ferrous (Fe^{2+}) forms are more soluble than the dietary ferric (Fe^{3+}) form, with twice the absorbability. The estimated absorption rate of the ferrous salts is 10-15%, with no difference found in absorbability between the different formulations. Note, though, that the different ferrous salts contain differing quantities of elemental iron per tablet. The equivalent of 60 mg of elemental iron is 300 mg ferrous sulphate heptahydrate, 180 mg ferrous fumarate or 500 mg of

ferrous gluconate. Oral solutions of ferrous iron salts are available for use in children, but can cause black staining of teeth.

For the treatment of iron-deficiency anaemia in adults, the recommended daily dose of elemental iron is in the range of 150 to 200 mg/day; for children the dose is 3-6 mg iron/kg body weight/day. Because absorption is so poor, thrice-daily dosing has traditionally been recommended to reverse iron-deficiency. For example, a single ferrous sulphate 300mg tablet contains 60mg of elemental iron, so thrice-daily dosing provides 180mg of elemental iron per day, well within the recommended daily range of 150-200mg for iron-deficient patients. If the absorption rate is 10%, then after a month of therapy about 500mg of bioavailable iron should have accumulated, which can produce 500mL of packed red blood cells and a haemoglobin increase of 2g/dL.

An understanding of how hepcidin regulates systemic iron absorption caused some researchers to rethink oral iron dosing. In 2015, Moretti et al. published their findings: A large oral dose of iron taken in the morning causes an increase in the plasma iron level. This stimulates an increase in hepcidin, which then interferes with the absorption of an iron dose taken later in the day. The suppression of iron absorption could last as long as 48 hours. They concluded that providing lower dosages and avoiding twice-daily dosing actually maximizes fractional iron absorption, and their results support supplementation with 40-80 mg of iron taken every other day.

Side effects of oral iron therapy, namely constipation, diarrhoea, black stools, heartburn, nausea, and epigastric pain, affect 35-59% of patients and limit compliance. The upper GI side effects, such as nausea and epigastric pain, are more dose-dependent and can be managed with lower or less frequent dosing initially, while lower GI effects such as altered bowel habits, are less related to dosing. There is no difference in GI side-effects between equivalent dosages of the different ferrous salt preparations. Taking iron with meals to minimize GI upset reduces absorption by almost 50% (phytates in cereal, tannins in tea, and foods rich in calcium hinder absorption).

Enteric-coated formulations of ferrous salts have been developed to decrease the prevalence of GI upset and reduce the dosing schedule. The problem with these is that the iron may not be absorbed in the duodenum. The estimated bioavailability of the enteric-coated preparations is 30% of the regular oral preparations.
A polysaccharide-iron complex formulation has been designed to minimize GI upset via delayed iron release in the intestines. The combination of ferric iron and low molecular weight polysaccharide contains 150mg elemental iron. Intestinal absorption is significantly less than that of equivalent dosages of ferrous salts, though GI side-effects are much less severe.
Carbonyl iron is used as a substitute for ferrous sulphate. It has a slower release of iron as it relies on gastric acid for its solubilisation and is more expensive than ferrous sulphate. The slower release affords the agent greater safety if ingested by children. On a milligram-for-milligram basis, it is 70% as efficacious as ferrous sulphate. It is claimed to cause fewer GI effects.
Co-administration with vitamin C: Ascorbic acid theoretically improves absorption by reducing iron to a ferrous Fe^{2+} state to optimize its solubility. Increasing doses of vitamin C exhibit a dose-dependent response in iron absorption during concomitant administration in healthy volunteers. Co-administration of 500mg vitamin C results in a 48% increase in the absorption of 30mg of elemental iron.

Intravenous iron therapy:

Treatment with IV iron presents several advantages over oral iron such as a faster and higher increase in Hb levels and replenishment of body iron stores. There are extended **indications** for intravenous iron compared with oral iron:

MAIN CLINICAL INDICATIONS FOR INTRAVENOUS IRON TREATMENT	
1. Intolerance of oral iron or non-compliance with an oral iron regimen	
2. In acquired or hereditary decreased intestinal iron absorption	Intestinal malabsorption syndromes, inflammatory bowel disease, after gastrectomy/bariatric surgery, iron-resistant iron-deficiency anaemia
3. In cases with severe iron-deficiency anaemia because of continuous or uncontrolled blood loss and/or because of increased iron needs	Hereditary haemorrhagic telangiectasias, angiodysplasia due to other causes, pregnancy (after 14 weeks), postpartum anaemia, patients scheduled for surgery
4. In cases of functional iron deficiency* particularly when an erythropoiesis-stimulating agent is being used	Anaemia of chronic kidney disease, inflammatory diseases, anaemia of cancer
5. Other circumstances	Autologous blood donation for elective surgery, Jehovah's Witnesses

* Functional iron deficiency is characterized by the presence of adequate iron stores as defined by conventional criteria but an inability to sufficiently mobilize this iron, particularly when erythropoiesis is stimulated by an erythropoiesis-stimulating agent. These patients, who often have inflammatory disease, respond faster & more completely to high doses of IV iron for treatment of anaemia.
But, what is the underlying mechanism? In these patients, a small proportion of the infused iron is delivered in the ferric form into the plasma and taken up by transferrin. Most of the administered iron dose is taken up by the macrophages. The iron overload of the macrophages in the RES may cause a 'bypass' of the hepcidin-induced block, allowing over-expression of ferroportin and allowing a flow to the bone marrow, transported by transferrin, to sustain erythropoiesis. In addition, in autoimmune diseases, macrophage iron loading may inhibit pro-inflammatory immune effector pathways, thus reducing disease activity (anti-inflammatory effect).

As free iron may lead to the production of hydroxyl radicals with potential toxicity to tissues, iron deficiency should be confirmed by ferritin levels before use of parenteral preparations.
Contraindications include a history of anaphylaxis or reactions to parenteral iron therapy, first trimester of pregnancy, active acute or chronic infection and chronic liver disease

Intravenous iron classification

Most IV iron formulations are colloids, comprising an iron-oxyhydroxide core (Fe3+) stabilized by a carbohydrate outer shell that maintains iron in a colloid form and controls its release. Formulations differ in iron core size and in the type and density of the surrounding carbohydrate shell. The stronger the iron complex, the slower the rate of iron release. IV iron complexes can generally be classified as labile or robust (kinetic variability) and as weak or strong (thermodynamic variability) with all possible intermediates. Each iron-carbohydrate complex enters the RES macrophages of the liver, spleen, and bone marrow where the shell is broken down and the iron is released from the complex. The released iron is either transported into storage pools or is transported via plasma transferrin for its incorporation into haemoglobin.

Older formulations

1) *Ferric hydroxide*: first iron compound for parenteral use; introduced early in the 20th century; lacked a carbohydrate shell → immediate iron release and severe toxic reactions; recommended only in extraordinary circumstances.
2) *Imferon*®: first high-molecular-weight iron dextran [HMW-ID] for intramuscular and IV use; introduced in 1954. HMW-ID consists of an iron oxyhydroxide core, which is surrounded by a carbohydrate shell made of polymers of dextran. This carbohydrate shell controls the release of free iron from the complex and also limits the total dose that can be given at any one administration. The bioavailability of iron occurs via uptake of iron dextran particles into the reticuloendothelial system (RES) with subsequent breakdown. High incidence of serious adverse events, especially the well-known dextran-induced anaphylactic reactions, led to its recommendation only when extreme clinical conditions were present and other options unavailable; removed from the market in 1991.

In 1989, recombinant human erythropoietin was introduced for clinical use, and it was recognized that absolute or functional iron deficiency was the commonest cause for erythropoietin failure in chronic kidney disease patients. IV iron therapy was shown to play an essential role in achieving and maintaining target Hb levels in CKD patients. This provided the impetus to develop better IV iron formulations.

3) *INFeD*®: low-molecular-weight iron dextran (LMW-ID); approved by FDA for clinical use in 1992; can be administered as an IV bolus or total dose infusion (TDI) with doses up to 1000 mg. Much lower risk of adverse events compared with the now-withdrawn HMW-ID compounds.
4) *Dexferrum*®: a HMW-ID, approved by FDA in 1996; administered as an IV bolus or total dose infusion (TDI) with doses up to 1000 mg; required a test dose and had a black box warning; no longer available in Europe.
5) *Ferrlecit*®: sodium ferric gluconate in sucrose (FG), after having been available in Europe for many years, was introduced into the American market in 1999 as a safer alternative to iron dextran with a lower risk of severe hypersensitivity reactions

6) *Iron sucrose* (IS) (Venofer®) was approved in the US in November 2000, and there is more published literature about this drug than any other IV iron preparation. IS can be safely administered as a 15-30 minute infusion in doses of 200-300 mg; the maximum weekly dose should not exceed 600 mg. If higher-than-recommended doses are not infused, AEs are rarely observed. The main disadvantage of IS is the need for multiple infusions as the maximum weekly dose should not exceed 600 mg (200 mg IV, 1-3 times/week).

The incidence of serious life-threatening anaphylaxis with IS is 0.002% *vs* 0.6-2.3% and 0.04% with HMW-ID and FG, respectively. Black box warnings do not appear in the directions for use of either FG or IS and a test dose is not required.

Newer formulations

Three new IV iron compounds have been released for clinical use in patients with iron-deficiency anaemia in the last 10 years. All three of these new compounds have better safety profiles than the more traditional IV preparations, particularly because these products may be given more rapidly and in larger doses than their predecessors.

1) *Ferric carboxymaltose* (FCM): dextran-free parenteral iron product; first new agent approved for rapid and high-dose replenishment of depleted iron stores; iron complex that consists of a ferric hydroxide core stabilized by a carbohydrate shell- the design of the macromolecular ferric hydroxide carbohydrate complex allows controlled delivery of iron to the cells of the RES and subsequent delivery to the iron-binding proteins, ferritin and transferrin, with minimal risk of releasing large amounts of ionic iron into the serum. Stable, with very low immunogenic potential. Large doses (15 mg/kg; maximum of 1000 mg/infusion) may be administered in a single and rapid (15-minute) infusion without the need for a test dose.
2) *Ferumoxytol* (FeraHeme®): Approved by the FDA in 2009 for iron replenishment in adult CKD patients with iron-deficiency anaemia. It consists of a superparamagnetic iron oxide that is coated with a carbohydrate shell, which helps to isolate the bioactive iron from plasma components until the iron-carbohydrate complex enters the RES macrophages. It can be administered as a relatively large dose (max 510 mg) in a rapid (< 20 seconds) session without test dose requirement. The main adverse effects are serious hypersensitivity reactions (0.2%) and transient interference with the diagnostic ability of magnetic resonance imaging for up to 3 months.
3) *Iron isomaltoside 1000* (Monofer®): introduced into Europe in 2010; a non-branched, non-anaphylactic carbohydrate, structurally different from the branched polysaccharides used in iron dextran. Very low immunogenic potential and a very low content of free iron and can therefore be administered as a rapid high dose infusion of up to 2000 mg without a test dose, which offers the possibility of providing one-dose iron repletion. Monofer® does not have a colloid structure; rather, the linear oligosaccharide isomaltoside 1000 allows the formation of a matrix with interchanging iron and carbohydrate.

In Australia and New Zealand, *iron polymaltose* (Ferrum H® & Ferrosig®) is registered for IM and IV use. It has an iron content of 50 mg/mL, and can be administered as a total-dose infusion, with a maximum dose of 2500 mg in patients >35 kg.

A patient's total body iron deficit can be calculated using the *Ganzoni formula:*

total body iron deficit, or **cumulative iron dose (mg) =**
body weight* (kg) x (target Hb – actual Hb in g/L) x 0.24 + iron stores (mg)*****

*Use ideal body weight in overweight patients. If underweight, use actual body weight
**The factor 0.24= 0.0034 x 0.07 x 1,000: For this calculation the iron content of haemoglobin = 0.34%, blood volume = 7% of the bodyweight, and 1,000 is the conversion from g to mg.
***Iron stores:
<35 kg body weight: iron depot = 15 mg/kg body weight
≥35 kg body weight: iron depot = 500 mg
For example a 70 kg female with Hb 8 g/dL has an iron deficit of:
70 x (150 – 80) x 0.24 + 500 = 1676 mg i.e. approx. 1700 mg

	High-molecular-weight iron dextran	**Low-molecular-weight iron dextran**	**Iron gluconate**	**Iron sucrose**	**Ferric carboxymaltose**	**Ferumoxytol**	**Iron isomaltoside 1000**
Brand name	Dexferrum ®	INFeD ®	Ferrlecit ®	Venofer ®	Ferinject ® Injectafer ®	FeraHeme ®	Monofer ®
Carbohydrate shell	Dextran (branched polysaccharide)	Dextran (branched polysaccharide)	Gluconate (monosaccharide)	Sucrose (disaccharide)	Carboxymaltose (branched polysaccharide)	Polyglucose sorbitol carboxymethylether	Isomaltoside (linear oligo-saccharide)
Complex type	Type I Robust & strong	Type I Robust & strong	Type III Labile & weak	Type II Semi-robust & moderately strong	Type I Robust & strong	Type I Robust & strong	Type I Robust & strong
Molecular weight (kDa)	265	165	289-440	30-60	150	750	150
Initial V_D (L)	3.5	3.5	6	3.4	3.5	3.16	3.4
Plasma half-life (h)	60	20	1	6	16	15	20
Labile iron release	-	-	+++	+-	-	-	-
Direct iron donation to transferrin (% injected dose)	1-2	1-2	5-6	4-5	1-2	<1	<1
Iron content (mg/mL)	50	50	12.5	20	50	30	100
Maximum single dose (mg)	20mg/kg	20mg/kg	125	200-300	15mg/kg (max 1000mg in one infusion)	510	20mg/kg
Total-dose infusion possible	Yes	Yes	No	No	No	No	Yes
Premedication	TDI only	No	No	No	No	No	No
Test dose required	Yes	Yes	No	No	No	No	No
Black box warning	Yes	Yes	No	No	N/A	No	N/A
Preservative	None	None	Benzyl alcohol	None	None	None	None
Life-threatening ADEs (x 10^6 doses)	113	3.3	0.9	0.6	?		?

Potential negative effects of IV iron, of concern but not all fully clarified:

1) *acute hypersensitivity reactions*: much rarer with decreasing use of HMW-ID, but can be life-threatening. The two likeliest mechanisms are immunological IgE-mediated responses, for example, to the dextran component of IV iron preparations containing this molecule, and complement activation-related pseudo-allergy (CARPA).
2) *vasoactive reactions*, which are due to the appearance of nontransferrin-bound or free (labile) iron in the circulation, include a drop in blood pressure, acute edema of extremities, and acute onset of diarrhoea when large IV iron doses are administered rapidly
3) If iron release exceeds binding capacity or if transferrin is oversaturated, toxic unbound iron results in ↑ *oxidative stress & free radical formation*—believed to lead to coronary artery inflammation, with atherosclerosis development and *long-term CV risk*. Iron may also cause LDL oxidation →coronary artery damage. MI occurring with the use of IV iron (possibly caused by ↑ serum ferritin levels) has been reported
4) *promotion of tumour growth* in patients with cancer
5) *endothelial dysfunction*
6) *renal tubular damage*- iron sucrose
7) theoretical risk of *infection*: elemental iron is an essential growth factor for bacteria, with many species expressing iron transport proteins that compete with transferrin; available evidence suggests no increased risk of sepsis, but some practitioners withhold IV iron if acute infection
8) *iron overload*: ↑ risk for liver disease (cirrhosis, cancer), MI or CCF, DM, osteoarthritis, osteoporosis, metabolic syndrome, hypothyroidism, hypogonadism and in some cases premature death. Iron overload can accelerate neurodegenerative diseases e.g. Alzheimer's, early-onset Parkinson's, Huntington's, epilepsy and multiple sclerosis.

References: Moretti D et al. *Oral iron supplements increase hepcidin and decrease iron absorption from daily or twice-daily doses in iron-depleted young women.* Blood 2015; 126(17):1981-1989
Further reading: Anderson G & McLaren G ed. *Iron Physiology & Pathophysiology in Humans.* Humana Press, 2012.

Nitric Oxide

Dr Caroline Simons

UCT Dept of Anaesthesia & Perioperative Medicine

Nitric Oxide (NO) is used by the body as a signaling molecule. It was previously known as EDRF (endothelial derived relaxing factor) and was named the “Molecule of the Year” by Science magazine in 1992 as the importance of the role of NO in the mammalian body became apparent. In 1998 Furchgott, Ignore and Murad won the Nobel Prize for elucidating the role of NO in the cardiovascular system.

Outside of the human body it is produced as the byproduct of industrial processes, especially by internal combustion engines, and by cigarettes. It is a greenhouse gas and is classified by many countries as an extremely hazardous gas.

Chemistry

Nitric oxide is a colorless gas at standard temperature and pressure. It has an *unpaired* electron on the nitrogen atom which makes it highly reactive. Common reactions include:

$2NO + O_2 = 2NO_2$ which is nitrogen dioxide and causes nausea, headaches and impaired immune and respiratory function.

$4NO + O_2 + H_2O = 4HNO_2$ (nitrous acid)

NO is naturally produced by lightning strikes. Since the heat generated is over 2000 degrees Celsius this reaction does not need a catalyst, ie $N_2 + O_2 = 2NO$

Commercially NO is prepared by the oxidation of ammonia at 850 degrees Celsius with platinum as a catalyst. The reaction is $4NH_3 + 5O_2 = 4NO + 6H_2O$

In the laboratory NO is produced by the reduction of dilute nitric acid with copper. $8NO_3 + 3CU = 3Cu(NO_3)_2 + 4H_2O + 2NO$

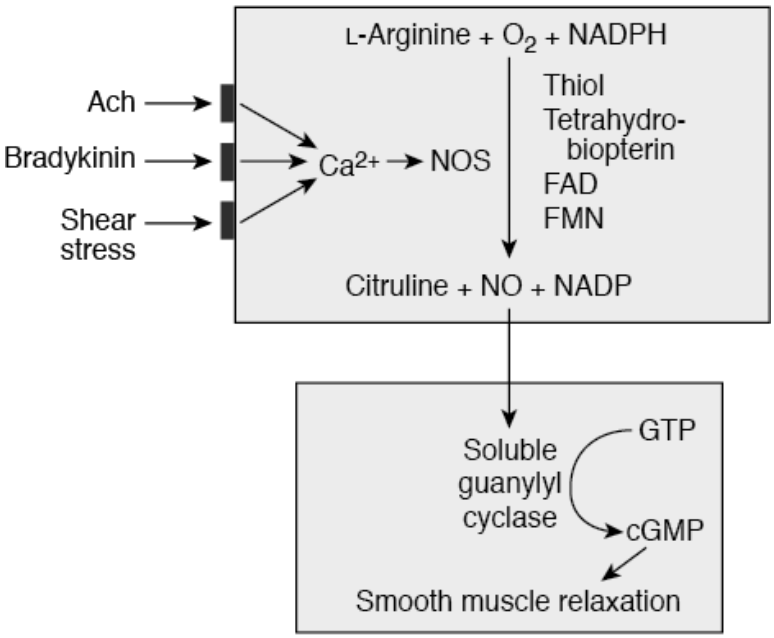
Medical grade NO is produced under carefully controlled conditions, diluted with nitrogen and stored in the absence of oxygen.

Levels are measured by chemoluminescent and electrochemical methods. Chemoluminescence is regarded as the gold standard but electrochemical is also accurate.

Physiology

NO is generated the reaction of L-arginine with oxygen and NADPH. Citrulline, NO and NADP are generated as end products. The reaction is catalyzed by NOS (nitric oxide synthase) which uses Thiol, tetrahydrobiopterin (TH4), FAD (flavin adenine dinucleotide), FMN (flavin mononucleotide) as requisite cofactors.

Nitric Oxide is a gaseous molecule and does not require channels or receptors. It diffuses across cell membranes to activate soluble guanylyl cyclase (sGC) which in turn will catalyze the formation of cGMP (cyclic guanosine monophosphate) from GTP (guanosine 5' triphosphate). cGMP can now activate smooth muscle relaxation.



Nitric oxide has a half-life of 2-6 seconds and is rapidly inactivated by blood. This makes its action highly specific to the area within which it is generated. Nitric oxide is mainly inactivated by binding to oxyhaemoglobin, forming methaemoglobin. Haemoglobin in turn will be regenerated forming NO_3^- (nitrate). A lesser proportion of NO will bind to deoxyhaemoglobin forming nitrosohaemoglobin or will dissolve with oxygen (O_2) to form NO_2^- (nitrite).

An alternative pathway to the formation of NO exists in the GIT. Nitrate rich vegetables (leafy greens) are metabolized by the nitrate reductase enzyme present in the GIT (mainly in the mouth) to nitrite (NO_2^-). The low pH in the stomach will then reduce nitrite to NO.

In the presence of hypoxia the nitrate-nitrite-nitric oxide pathway can also be activated intracellularly by deoxyhaemoglobin or deoxymyoglobin.

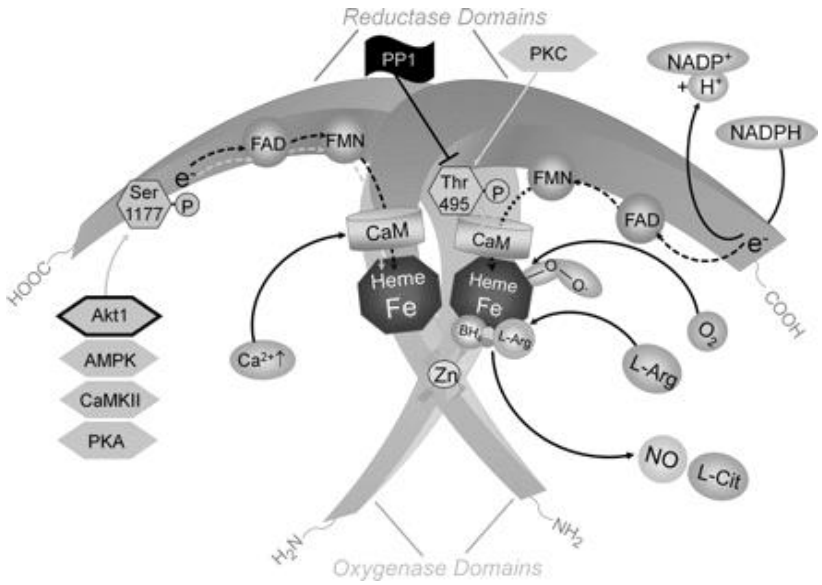
Primarily however the majority of the NO present in the body is catalyzed by NOS and therefore I will focus in the enzyme.

Nitric Oxide Synthase

Recommended reading: Nitric Oxide Synthases: Regulation and function by Forstermann and Sessa in European Heart Journal of April 2012

Three isoforms of the enzyme exist. NOS 1 is constitutive and present in neuronal cells. It is also known as nNOS. The expression of NOS2 is induced by factors such as bacterial endotoxins and cytokines. It is therefore known as iNOS, and when induced can be found in multiple cells, most notably in macrophages and neutrophils. NOS3 (eNOS) is constitutive and is present in multiple cells, but is well known for its function in endothelium and platelets. Constitutive NOS enzymes (1 and 3) are normally present in cells and are responsible for multiple baseline functions. They produce NO under normal physiological conditions in picomols. Induced NOS is produced in times of physiological stress and then larger amounts (nanomols) are produced. All NOS contain haem which is essential to their function. NOS also contains zinc but it is a structural element and does not catalyze reactions.

NOS enzymes are homodimers. NOS transfers electrons from NADPH via the flavins in FAD and FMN in the reductase domain to the haem in the oxygenase domain. The oxygenase domain binds BH4, O_2 and L-arginine. Now NOS will hydroxylate L-arginine. Secondly NOS oxidizes the N-hydroxy-L-arginine formed in the previous step to L-citrulline and NO.



Calmodulin binds to all isoforms of NOS and regulates their function. In nNOS and eNOS calmodulin binding occurs when there is an increase in intracellular calcium. This will then facilitate electron transfer, increasing levels of NO and cause smooth muscle relaxation. Calmodulin can bind to iNOS at low intracellular calcium levels.

NOS1

nNOS can be found in brain tissue, the spinal cord and some peripheral nerves. Nerves containing nNOS are known as nitrergic nerves. NO acts as a neurotransmitter in the CNS and autonomic nervous system. NO is involved in learning and memory by being involved in long term neuronal inhibition and potentiation of signals. In this way NO seems to help neurons “remember” previous signals. NO is not used by the CNS as an “acute” neurotransmitter. NO plays a role in the modulation of the state of arousal, pain perception, neurogenesis and apoptosis. NO is also present the medulla and hypothalamus, thereby playing a role in the central regulation of blood pressure. Furthermore nitrergic nerves innervate smooth muscle in the periphery, such as vascular smooth muscle, leading to vasodilation.

Pathyphysiologically abnormal NO signaling has been implicated in neurodegenerative conditions such as Alzheimers, Parkinsons and multiple sclerosis. In the presence of an influx of intracellular calcium (such as after a cerebrovascular incident) NMDA activation leads to cell death. NO participates by forming peroxynitrite ($OONO^-$) on reaction with the superoxide anion (O_2^-). Peroxynitrite is a potent oxidant and will cause cellular damage.

Nitrergic nerves also assist with the control in regional blood flow in the corpus cavernosum. eNOS is also present in penile endothelium. NO is therefore critical to penile erections.

nNOS can also be found in adrenal glands, kidney macula densa cells, pancreatic islet cells and epithelial cells of various organs.

NOS2

NOS2 is not normally found in cells. NOS2 expression is induced by substances such as cytokines and bacterial endotoxins. Corticosteroids can inhibit the formation of NOS. NOS2 is used by several cell types. In neutrophils and macrophages NO is formed which is toxic to some pathogens and tumour cells by causing direct DNA damage as well as degradation of the iron sulphur centers of enzymes critical for mitochondrial electron transport, cis-aconitase and ribonucleotide reductase. The high levels of NO produced can also damage surrounding healthy cells.

NOS2 expression can also be induced in multiple other cells such as myocytes and endothelial cells. In endothelial cells it has been implicated in the hypotension and capillary leak found in septic shock. In reperfusion injuries NO has been implicated in cellular damage via the formation of peroxynitrite ($OONO^-$) when NO reacts with superoxide. Peroxynitrite will harm the surrounding cells.

NOS3

As mentioned previously calmodulin regulates the activity of this enzyme. In addition heat shock protein 90 (hsp90) can activate eNOS and caveolin-1 is a tonic inhibitor of eNOS activity. Caveolin-1 can be displaced from eNOS by calmodulin and hsp90.

eNOS can also be activated by shear stress. This is achieved by the phosphorylation of the enzyme. This is achieved in several ways. Shear stress itself activates protein kinase A (PKA). Akt and AMPK (AMP activated protein kinase) are activated by insulin. Bradykinin activates Ca++/calmodulin dependent protein kinase II (CaMKII). Oestrogen and vascular endothelial growth factor (VEGF) also activate Akt. Enzyme phosphorylation produces increased electron flux in the reductase domain and increases the calcium sensitivity of eNOS.

Protein kinase c (PKC) can inhibit eNOS function by causing the phosphorylation of a different part of the enzyme.

NOS3 is constitutively present in multiple cell lines, including the following:

1) Endothelial cells

Here NO generated in the endothelium diffuses across to the vascular smooth muscle. By activating sGC, cGMP is formed. cGMP in turn activates protein kinase G which phosphorylates a number of proteins involved in smooth muscle relaxation. This will result in a drop in intracellular calcium and the inactivation of myosin light chain kinase.

In the pulmonary vasculature NO also inhibits hypoxic pulmonary vasoconstriction. NO will enhance flow in well ventilated areas which will improve the ventilation/ perfusion matching. Elevated levels of pulmonary NO play a role in attenuating hypoxia at high altitude.

In the vascular tree NO has anti-Inflammatory and anti-atherosclerotic effects

- Inhibits leucocyte adhesion and vascular inflammation (early stages of atherosclerosis)
- Prevents fibrous plaque formation by controlling smooth muscle proliferation (later stages of atherosclerosis)
- During oxidative stress NOS is “uncoupled” and will produce superoxide instead of NO. This is done by the oxidation of BH4, L-arginine depletion, accumulation of ADMA (asymmetric dimethyl arginine) which competes with L-arginine and S-glutathionylation of eNOS. Thereby patients with cardiovascular risk factors produce inadequate NO and instead produce reactive oxygen species.

NO also stimulates angiogenesis postnatally and after ischaemic events. NO is also thought to active endothelial progenitor cells in the bone marrow.

2) Platelets

NO increases permeability of the K+ pump which will hyperpolarize the cell membrane which inhibits contraction. In this way NO prevents platelet aggregation and adhesion. NO also prevents the release of platelet derived growth factors. These growth factors stimulate smooth muscle proliferation. eNOS can therefore play a role in preventing angiogenesis. NO production by eNOS is essential to adaptive vascular remodeling to chronic flow changes.

3) GIT

NO determines motility and modulates morphine induced constipation.

4) Kidneys

NO plays a role in sodium homeostasis in the kidney. NO also increases renal blood flow and GFR.

PHARMACOLOGY

Inhaled NO (iNO)

The lungs have high levels of NO. They arise from various cell types, especially endothelial cells and nitrergic nerves, but also from neutrophils, macrophages, epithelial cells, fibroblasts and smooth muscle cells. NO plays a critical role in the lungs. It is responsible for maintaining low pulmonary artery pressure via NO production by pulmonary eNOS. NO also opposes the pulmonary response to endogenous and exogenous vasoconstrictors, opposes hypoxic pulmonary vasoconstriction, controls pulmonary blood flow distribution and has an important immune function when expressed in neutrophils. NO may control bronchomotor tone via nitrergic nerves. In pulmonary hypertension decreased expression of eNOS is found.

iNO causes direct relaxation of the pulmonary vasculature therefore decreasing pulmonary vascular resistance, pulmonary artery pressure and RV afterload. As NO is subsequently rapidly deactivated by the circulating blood to form methaemoglobin, it has no effect on the systemic blood pressure. Thus it has no effect on coronary perfusion. By improving RV performance the LV filling pressure may increase. This is only problematic in patients with severe left ventricular dysfunction when pulmonary edema may result.

NO opposes hypoxic pulmonary vasoconstriction and controls pulmonary blood flow distribution, redirecting blood to well ventilated areas, thereby improving ventilation/ perfusion matching and decreasing right to left shunting of deoxygenated blood. iNO can improve oxygenation via these methods.

iNO may cause some bronchodilatory effects, thereby decreasing alveolar dead space. iNO can have anti-inflammatory (decreasing capillary leak) and antiproliferative effects. It will also decrease platelet aggregation and adhesion.

There have been concerns that NO may cause hypotension and bleeding (e.g. intraventricular haemorrhage in neonates), but this has been unfounded. 70% of the iNO dose will be excreted in the urine within 48hours as nitrate.

Side effects of iNO do exist though, namely: methaemoglobinaemia, pulmonary toxicity and rebound hypertension. Oxyhaemoglobin inactivates NO to form methaemoglobin. In adult patients the levels of methaemoglobin formed are not high enough to cause problems if the iNO dose remains less than 40ppm. In patients with a glucose 6 phosphate dehydrogenase deficiency however, methaemoglobinaemia levels may become high enough to impair oxygenation. It should also be noted that fetal Hb is more readily oxidized to methaemoglobin by NO.

In the presence of oxygen, nitrogen dioxide is formed ($2NO + O2 = 2NO2$). Nitrogen dioxide is a respiratory irritant, causing chest pain, bronchospasm and pulmonary edema. This can become problematic if high doses of iNO are administered, or a high FiO2 is used (most institutions will try and keep it under 60%).Neonates, the elderly, patients with respiratory and cardiac failure are more susceptible. Therefore it is imperative that the amount of nitrogen dioxide and NO in the circuit is monitored during iNO administration. NO on its own can also cause DNA damage to the lung when peroxynitrite is formed.

While administering iNO, endothelin 1 is upregulated and NOS is downregulated. If iNO is rapidly discontinued, rebound pulmonary hypertension and/or hypoxia will occur. Therefore iNO will always be weaned off slowly in a stepwise fashion. If the patient needs to be disconnected from the ventilator for transport or for ventilation by hand for incidents such as worsening hypoxia, the iNO must be continued.

iNO has a dosing range of 5-80ppm, but doses over 20ppm have been shown to have little additional benefit.

Other disadvantages of iNO are that is requires specialized expensive equipment, is expensive and cumbersome to administer, makes transporting the patient difficult, delivery system errors may occur, and it can pose an environmental hazard to hospital staff (headaches and respiratory irritant).

South Africa has set the OEL (occupational exposure limit) for NO at 25ppm and for nitrogen dioxide at 3ppm. At levels of 100ppm NO is regarded as immediately dangerous to life and health.

Uses of iNO:

a) Adults

ARDS (Acute Respiratory Distress Syndrome)
In ARDS iNO improves oxygenation and decreases shunting via aforementioned actions. In a small number of patients with a patent foramen ovale their pulmonary artery pressure can be high enough to shunt blood through the PFO. iNO can decrease the pulmonary artery pressure enough to reverse this.

iNO will improve oxygenation more if alveolar recruitment is improved. Therefore iNO is used with maneuvers to improve this, such as optimal PEEP, high frequency oscillatory ventilation or proning.

Initial studies have demonstrated the improvement in hypoxia but could not show an effect on mortality. They also showed an unexpected risk of AKI (acute kidney injury). The risk of AKI was initially put down to poor study design, but a Cochrane Database Review published in Anaesthesia in 2017, which included more recent studies, have confirmed this risk. They have concluded that iNO improves the PaO2 at 24 hours, but not at 48 or 72 hours. It also makes no difference in ventilator free days, duration of the ventilation, resolution of multiorgan failure, length of stay in ICU or hospital or quality of life. iNO also makes no difference in mortality. Furthermore the use of iNO significantly increases the risk of AKI. They found insufficient evidence to recommend the use of iNO in ARDS.

The mechanism of AKI is unclear. Some put it down to the formation of reactive nitrogen species, others to the increase in FiO2 required in ARDS. Both can produce a pro-inflammatory response with renal vasoconstriction.

Munshie and Adhikan point out in Critical Care of 2017 that it is difficult to design a good study in life threatening hypoxaemia in ARDS patients in ICU due to the multiple variables present (FiO2, ventilator strategies, different antibiotics, differing disease aetiology and imaging with contrast among others).

COPD
iNO will not improve gas exchange, and may worsen it.

Pulmonary Hypertension
iNO is used perioperatively for pulmonary hypertension. It decreases the need for postoperative ECMO. Cases in which it has proven useful include mitral valve surgery with significant pulmonary hypertension, heart transplant and LVAD (LV assist device) implantation. It can also be useful in adults with congenital cardiac disease associated with pulmonary hypertension. iNO has been used as salvage therapy in cases such as acute pulmonary embolism with right heart failure. There are varying responses to iNO in pulmonary hypertension due to different degrees of vascular remodeling and smooth muscle hypertrophy.

b) Paediatrics

Persistent Pulmonary Hypertension of the Newborn (moderate or severe)
This can be idiopathic/ primary or secondary due to diseases like meconium aspiration syndrome, pneumonia from group B strep or prematurity with respiratory distress syndrome. iNO reduces the need for ECMO. It is contraindicated in neonates dependent on right to left shunting of blood (certain congenital cardiac lesions).

Prematurity with Respiratory Distress Syndrome
It has been used in the past to improve oxygenation. It was also hoped the anti-inflammatory properties of NO along with its role in stimulating angiogenesis and activating endothelial progenitor cells would help to decrease the development of BPD (bronchopulmonary dysplasia) and neurodevelopmental injury. Studies have therefore looked at the routine and rescue use, both early and late. The Cochrane database review published in January 2017 has concluded that it is not effective as rescue for the critically ill premature infant. Early routine use does not prevent serious

brain injury or improve survival without BPD. Later use of iNO may be effective but the difference is small and requires further study. There were concerns that the effect of NO on platelets may increase intraventricular haemorrhage but this is not the case.

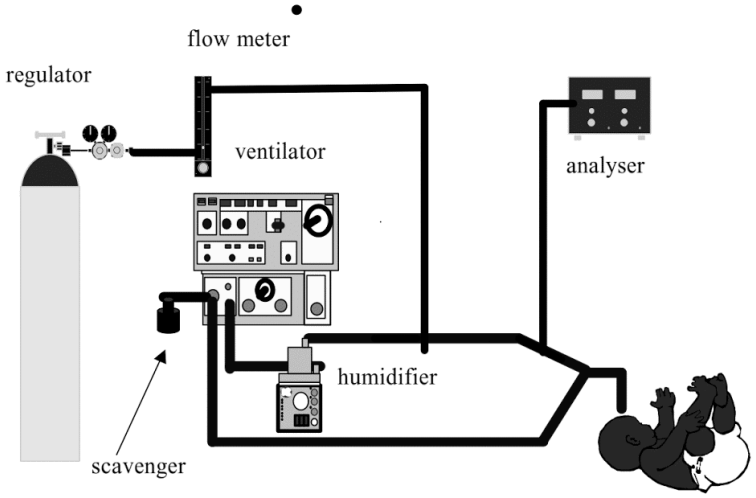
Congenital Diaphragmatic Hernia
NINOS trial demonstrated that although iNO was effective in improving oxygenation initially, it did not reduce the need for ECMO or decrease mortality.

Perioperative
iNO can be useful perioperatively as salvage therapy when there are problems with pulmonary hypertension, for example when coming off cardiac bypass with congenital cardiac lesions associated with pulmonary hypertension, such as AVSD.

c) Diagnostic

iNO has been used to diagnose reversibility with pulmonary hypertension.

Administration of iNO:



- **Cylinder**
  The cylinder contains medical grade NO with 1000ppm in nitrogen. It is under high pressure.
- **Pressure Gauges and Regulator**
  The regulator decreases the pressure from the cylinder. The first gauge demonstrates the pressure in the cylinder while the second one reflects the output pressure.
- **Flowmeters**
  Two flowmeters to determine the flow and therefore dose of iNO. The first one is for flow of 0-600ml and is used for larger adjustments of flow. The second one is for flow of 0-100ml and is used for smaller incremental adjustments of the iNO.
- **T-connector**
  Connects onto the inspiratory limb of the ventilator or oscillator circuit. Connect after the humidifier. It is important to connect it close to the patient in order to minimise formation of NO2. Do not connect closer than 20cm from the patient however, as then there is inadequate time for the NO to mix with the ventilator gases. Flush the delivery system before use to rid it of nitrogen dioxide/ nitric acid/ water or high oxygen levels. The ventilator or oscillator must have a constant flow rate to result in reliable levels of NO. If not the delivered dose of iNO is unpredictable. If the ventilator does not use a constant flow rate specialised injector systems need to be used.
- **Water trap**
  When combined with H2O nitrogen dioxide forms nitric acid which is also toxic.
- **Monitoring**
  Sidestream electrochemical monitoring. Calibrate before use. It is esssential to monitor both NO and NO2 levels as both can be toxic.
- **Scavenging filter**
  Removes NO and NO2 and ensures safe working environment for staff. When in use environmental levels of NO2 are usually less than 1ppm.

DRUGS ACTING INDIRECTLY ON THE NITRIC OXIDE PATHWAY

NO donors

Drugs such as nitroglycerin, isosorbide dinitrate and sodium nitroprusside. Please see your pharmacology textbook.

Phosphodiesterase 5 inhibitors (Sildenafil, Tadalafil, Vardenafil)

Phosphodiesterases are the enzymes responsible for the breakdown of cGMP to GMP. By inhibiting this enzyme these drugs raise the amount of cGMP in the cell and enhance the effect of NO. Phosphodiesterase 5 is the predominant isoform of the phosphodiesterase enzyme in penile tissue. These drugs are well known for their effect in enhancing penile erection (nNOS and eNOS present in the corpus cavernosum).There must be residual nNOS for the drugs to have any effect.

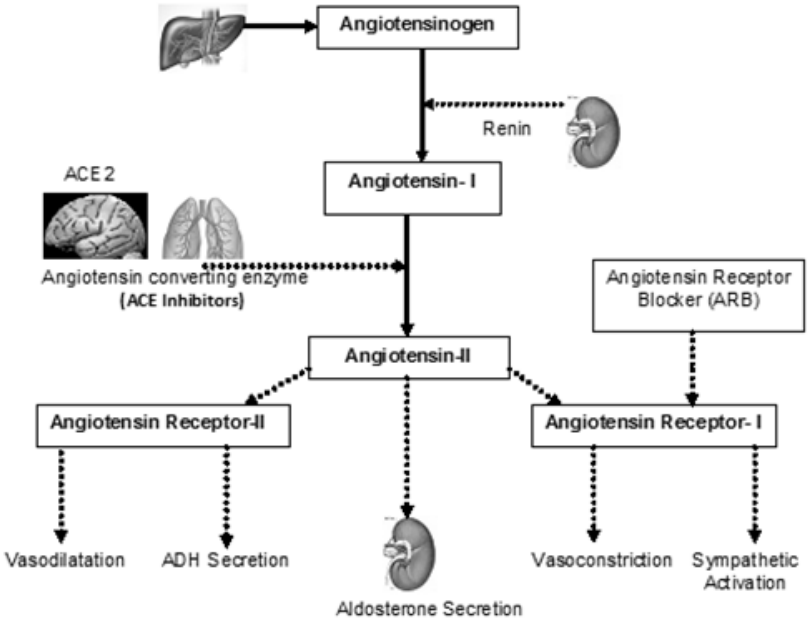
Phosphodiesterase 5 is also well represented in pulmonary arteries, and sildenafil and tadalafil are used in the treatment of pulmonary arterial hypertension. Side effects include headache, flushing, dyspepsia, nasal congestion and visual disturbances. More severe effects include priapism, hypotension (especially when administered with other drugs acting on the NO pathway), myocardial infarctions, arrhythmias, cerebrovascular incidents, increased intraocular pressure and sudden hearing loss.

These drugs can also potentiate and prolong the effect of iNO.

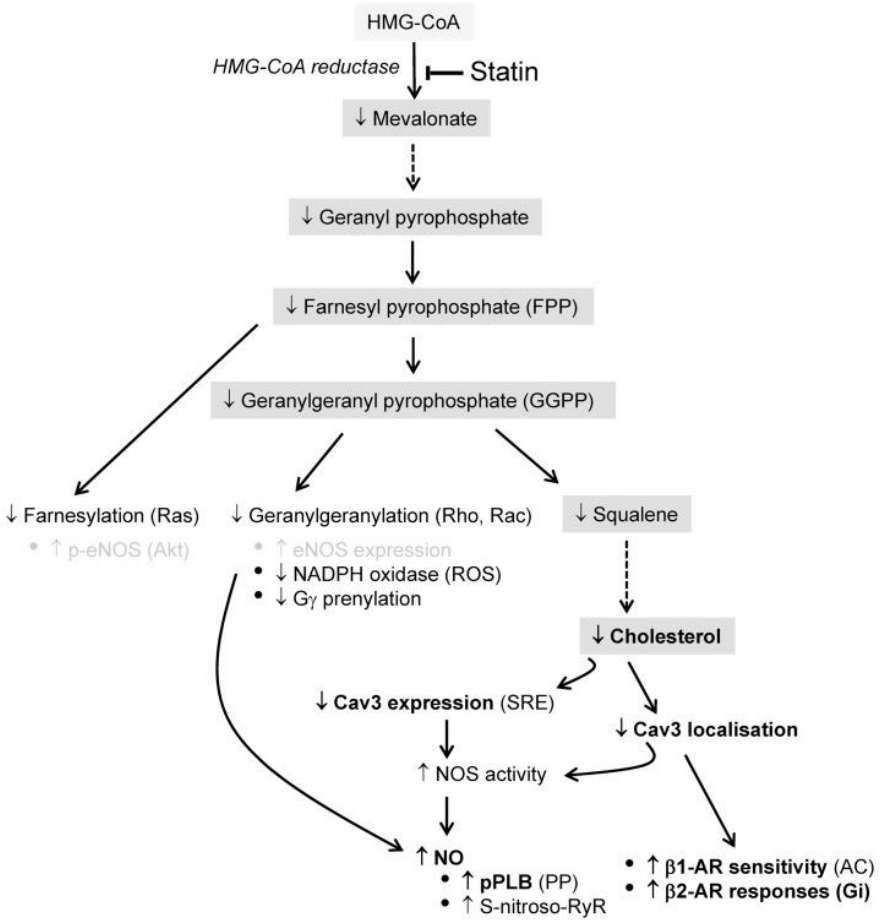
Pleiotropic Drugs

These drugs prevent eNOS uncoupling and in this way prevent endothelial dysfunction. They are regarded as cardioprotective as they prevent endothelial dysfunction. These include drugs that act on the renin- angiotensin- aldosterone pathway and statins.

ACE Inhibitors and Angiotensin Receptor Blockers
They lower blood pressure via their action on the renin/angiotensin/aldosterone pathway. Their pleiotropic effect is not their main mechanism of action but may be very important. These drugs are allosteric enhancer of the bradykinin receptors. This results in phosphorylation of eNOS and will then increase NO levels. AT1 receptors are upregulated by increased levels of LDL. Angiotensin 2 activates NADPH oxidases via AT1 receptors. Therefore ACE inhibitors and ARB's can prevent activation of NADPH oxidase and uncoupling of eNOS. These drugs can also increase the activity of superoxide dismutase (SOD3) which will scavenge superoxide. They may increase expression of GTP cyclohydrolase1 which is the rate limiting enzyme for BH4 synthesis, thereby increasing BH4 levels.



Statins
Their pleiotropic effects include increasing the expression of eNOS, enhancing the activity of eNOS by decreasing caveolin levels and activation of the phosphatidylinositol 3-kinase/ Akt pathway. Statins also decrease expression and activity of NADPH oxidase by preventing isoprenylation of p21Rac. Simvastatin more than doubles superoxide dismutase activity. Statins increases BH4 levels by increasing GTP cyclohydrolase1 mRNA expression in endothelial cells. This means that statins can assist in plaque stabilization, decrease thrombogenic responses, improve endothelial function and inhibit oxidative stress and inflammation.



As demonstrated above statins lower cholesterol levels by reversible inhibition of HMG-CoA reductase. This enzyme is the rate limiting step in cholesterol synthesis by the liver. Statins also inhibit the generation of isoprenoid intermediates such as FFP and GGPP. These isoprenoids can form lipid attachments on various proteins, a process known as isoprenylation. These include G-proteins and GTP binding proteins such as Rac,Ras and Rho. Statins therefore inhibit isoprenylation of these proteins, which results in accumulation of inactive Rac, Ras and Rho. This plays a large part in increasing NO levels.

Corticosteroids

Glucocorticoids will bind to the glucocorticoid receptor (GR) which will stimulate phosphatidylinositol 3-kinase and protein kinase Akt. This will cause activation of eNOS and NO dependent vasorelaxation. This is a non nuclear action and is a cardiovascular protective effect of high dose corticosteroids.
On the other hand, the nuclear actions of corticosteroids inhibit the expression of iNOS which is anti-inflammatory and may assist with the treatment of capillary leak and hypotension in septic shock.

sGC Stimulators

Riociguat is a drug available in the US which stimulates sGC. This will increase cGMP resulting in vasodilation. It is licensed for use for inoperable or persistent recurrent chronic thromboembolic pulmonary hypertension postoperatively, the treatment of pulmonary artery hypertension, and is being

investigated for use in Raynaud's disease and systemic sclerosis. It is contraindicated in pregnancy and pulmonary arterial hypertension from idiopathic interstitial pneumonia. Side effects include hypotension, bleeding, headaches and GI disorders. Drug interactions are with the nitrates and PDE5 inhibitors. Smoking will decrease Riociguat's levels.

Cinaciguat is an experimental drug for the use of acute decompensated heart failure.

OTHER USES

Exhaled NO can be used as an "inflammometer" in asthma. It has also been used to predict kids at risk of perioperative adverse respiratory events. Ramgolam et al in Australia studied this and concluded that an accurate history assessing risk factors remains the most appropriate tool to diagnose children at risk for adverse respiratory events.

References

1. Forstermann and Sessa. Nitric oxide synthases: regulation and function. European Heart Journal. 2012/ 33(7)/ 829-837
2. Ganong. Review of Medical Physiology. Twenty first edition.
3. The nursing staff teaching pack from the INOVO trial and the information pack from NOxBOX
4. Hafezi-Moghadam et al. Acute cardiovascular protective effects of corticosteroids are mediated by non-transcriptional activation of endothelial nitric oxide synthase. Nature Medicine 2002 May 8(5): 473-479
5. Anaesthesia UK site – Nitric Oxide
6. Troncy et al. Review Article: Inhaled Nitric Oxide: clinical applicaions, indications, and toxicology. Can J Anaesth 1997/ 44:9/ 973-988
7. Kam and Govender. Nitric oxide: basic science and applications. Anaesthesia 1994/ 49/ 515-521
8. Munshi and Adhikari. Inhaled nitric oxide and acute kidney injury: new insights from observational data. Critical Care 2017/ 21/ 83
9. Bhatraju et al. Inhaled nitric oxide: Current clinical concepts. Nitric Oxide 2015/ Aug 31/ 50/ 114-128
10. Thunberg et al. Inhaled therapy for the management of perioperative pulmonary hypertension. Annals of Cardiac Anaesthesia. 2015/ 18(3)/ 394-402
11. Ichinose et al. Inhaled Nitric Oxide: A selective pulmonary vasodilator. Current uses and therapeutic potential. Circulation 2004/ 109/ 3106-3111
12. Griffiths and Evans. Inhaled nitric oxide therapy in adults. NEJM. 2005/ 353/ 2683-2695
13. Bernasconi and Beghetti. Inhaled nitric oxide applications in paediatric practice. Images in Paediatric Cardiology. 2002/ 4(1)/ 4-29
14. Barrington et al. Inhaled nitric oxide for respiratory failure in preterm infants. Cochrane Database Syst Rev Jan 2017
15. Petit et al. The Pathophysiology of nitrogen dioxide during inhaled nitric oxide therapy. ASAIO J. Jan/Feb 63(1)7-13
16. Yang et al. Inhaled nitric oxide in preterm infants: An updated metananalysis. J Res Med Sci Jun 2016/ 14/ 21:41
17. Ramgolam et al. Prediction of perioperative adverse respiratory events in children: the role of exhaled nitric oxide. Anaesthesia 2015 Oct 70 (10)/ 1160-4
18. Karam et al. The effect of inhaled nitric oxide in acute respiratory distress syndrome in children and adults: a Cochrane Systematic Review with trial sequential analysis. Anaesthesia 2017 Jan 72(1) 106-117
19. Kinsella et al. Early inhaled nitric oxide therapy in premature newborns with respiratory failure. NEJM 2006 Jul 27/ 355(4)/ 354-64
20. Ballard et al. Inhaled nitric oxide in preterm infants undergoing mechanical ventilation. NEJM 2006 Jul 27/ 355(4)/ 343-53

Pharmacological Aspects of Serotonin for the Anaesthetist

Dr Kamal Bhagwan

UCT Dept of Anaesthesia & Perioperative Medicine

Serotonin (5HT) Physiology

In the **CNS** serotonergic neurons are found primarily in the **brainstem** from midbrain to medulla. The rostral end of this system assists in the regulation of wakefulness, affective behaviour, food intake, thermoregulation, migraine, emesis and sexual behaviour. The neurons in the raphe of the lower pons and medulla participate in the regulation of nociception and motor tone.

In the **PNS**, serotonin is produced primarily by intestinal enterochromaffin cells (95% of the total amount of serotonin of the body). Here it regulates gastrointestinal motility, vascular tone, uterine contraction, and bronchoconstriction. It also functions to promote platelet aggregation.
Serotonin, released by presynaptic serotonergic neurons in synaptic clefts, activates specific receptors and is partly reuptaken by presynaptic neurons

There are seven distinct serotonin receptors (5-HT1to 5-HT7 Rc) with further subtypes. Most are **G-protein** coupled and via adenyl cyclase or phospholipase C produce their effect.

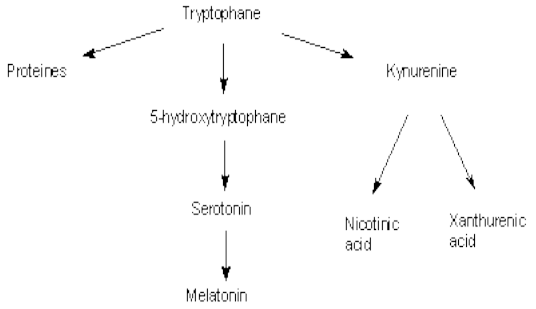
5-HT3 Rc is unique in that it belongs to the Cys-loop superfamily of ligand-gated **ion channels.** The activating inward current is cation-selective which is largely with sodium and potassium ions.

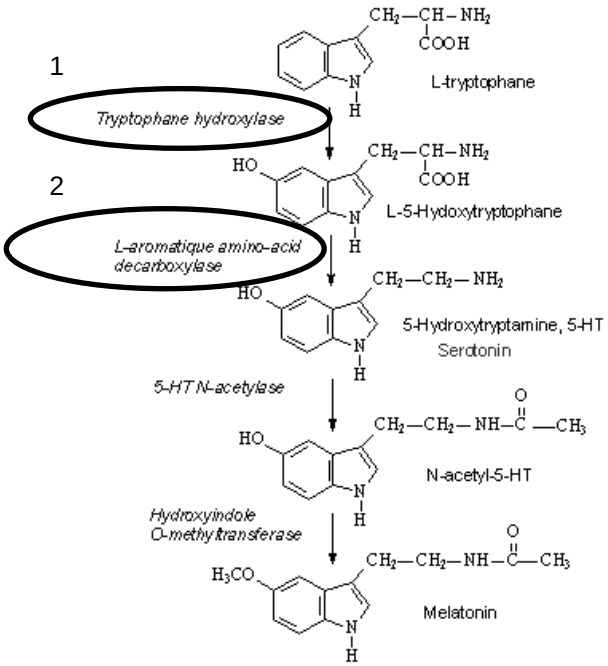
Biosynthesis

In the brain, serotonin biosynthesis depends on the quantity of tryptophan that crosses the blood-brain barrier.

Circadian variations determine the concentration of melatonin that increases during the night and decreases during the day.

5-HT is produced by 1. *hydroxylation* and 2. *decarboxylation* of L-tryptophan in nerve terminals and is stored in synaptic vesicles



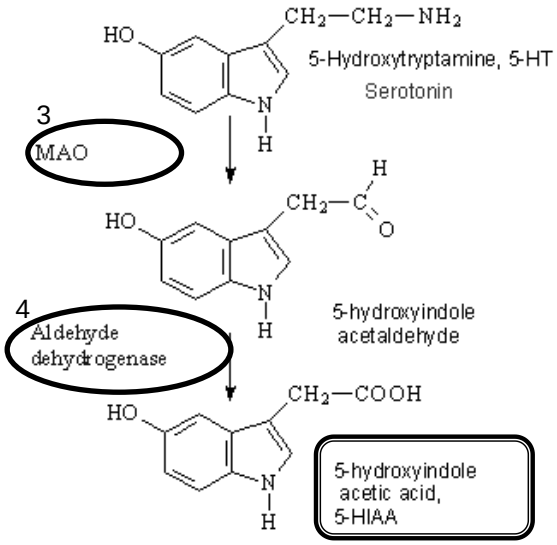


Inactivation

Serotonin is converted into inactive molecules by biotransformations:

3. *Oxidative deamination* of the lateral amino chain by monoamine oxidase, leading to 5-hydroxy-indol-acetaldehyde which is then oxidized into 5-hydroxy-indol-acetic acid (5-HIAA) found in urine

4. *Conjugation* by glucuronic acid or sulfate of the hydroxyl group OH in 5-position.



Exploration of drugs that interact with various 5HT receptors (Rc) follows:

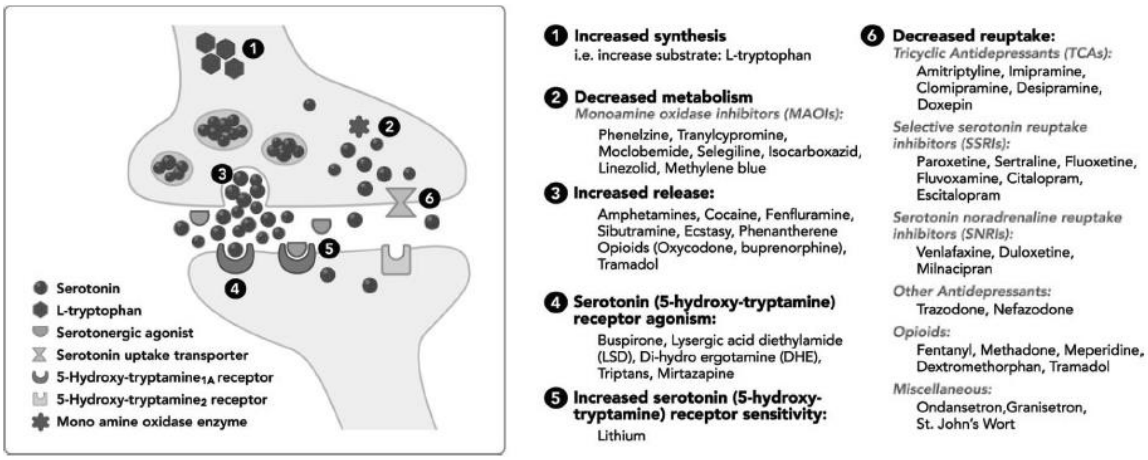


Fig. 1. Increase intrasynaptic serotonin levels: Mechanisms and associated serotonergic agents.

Anesthesiology 2011; 115:1291–8 1293 Rastogi *et al.*

5-HT1 Agonists

Triptans are utilized in the treatment of **migraine attacks** and are effective in the treating the acute attack. Examples of these drugs are sumatriptan, zolmitriptan, naratriptan, eletriptan, almotriptan and rizatriptan.

Sumatriptan

- Most studied of the triptans.
- Agonist of 5-HT1, 5-HT1D Rc's, and to a lesser degree of 5-HT1B
- Causes vasoconstriction of cerebral vessels, in particular of carotid arteriovenous anastomose
- Indicated for migraine and cluster headache attacks but not for their long- term prevention.

- 70% effective but short half-life and potential for coronary vasospasm and arrhythmias limits its use

Newer agents

- longer half-lives and fewer side effects.
- Selective carotid vasoconstriction via 5-HT1B receptors and by pre-synaptic inhibition of the trigeminovascular inflammatory response via 5-HT1D/1F receptors.

Anaesthetic implications

- Elucidate individual patients' triggers, symptom patterns, and preferred therapies.
- Continue preventive medication.
- Minimise variations in arterial blood pressure, temperature, and arterial CO_2.

Ergotamine

- Partial agonist at 5HT1 and alpha receptors → vasoconstriction and inhibition of trigeminal nerve transmission.
- Duration of action is 24 hrs
- Wide side-effect profile limits its clinical use - coronary vasoconstriction, severe nausea and vomiting, uterine contraction and foetal damage.

5-HT3 Antagonists: Anti-emetic Agents (-setrons)

The primary function of these drugs is to inhibit vomiting induced by antineoplastic drugs, such as cisplatin and doxorubicin. These drugs are cytotoxic and release serotonin from the digestive tract. Poor efficacy in the treatment of motion sickness as the vestibulo-ocular component of the afferents to the area postrema may be an important factor.

Other examples of drug-induced vomiting are attributed to opiates, radiotherapy and certain antibiotics. General anaesthetics can produce emesis by both peripheral as well as central mechanisms.

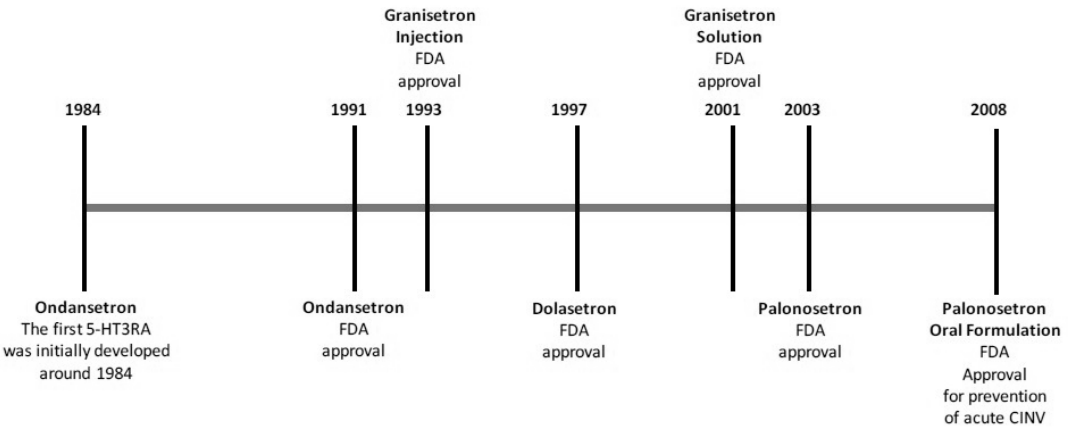
Mechanisms of action:

1. blocking 5-HT Rc's in the **Chemoreceptor Trigger Zone** within the **Area Postrema** and the **Nucleus Tractus Solitarius**

2. blocking **peripheral afferent vagal impulses** originating from 5-HT Rc's in **GIT mucosa** and inhibit splanchnic afferent nerve response to painful distension

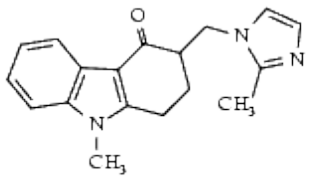
5-HT3 Rc antagonist drugs: ondansetron, granisetron, tropisetron and dolasetron.

Adverse effects: headache, asthenia, drowsiness, arterial hypertension or hypotension, bradycardia, hiccup, constipation.



Ondansetron

- Favourably rated in the treatment of postoperative nausea
- More efficacious than metoclopramide and/or droperidol.
- Very wide therapeutic "window" with only rarely occurring side effects
-



Ondansétron

Categorisation of chemical structures of the first generation 5-HT3 receptor antagonists

- Carbazole derivatives - Ondansetron
- Indazoles - Granisetron
- Indoles - Dolasetron

Annals of Palliative Medicine, Vol 1, No 2 July 2012 117

Table 1 5-HT3 receptor antagonists

Agent	Trade name	$T_{1/2}$ (h)	Dose	Metabolism (primary pathway)	Metabolism (secondary pathway)	Receptor subtypes
Ondonsetron	Zofran	3.9	0.15 mg/kg	CYP3A4	CYP1A2, CYP2D6, CYP2E1	5-HT3B, 5-HT1B, 5-HT-1C, α-1, MOR
Granisetron	Kytril	9-11.6	10 mcg/kg	CYP3A4	----------	5-HT3A, 5-HT3B, 5-HT3C
Dolasetron	Anzemet	7-9	0.6-3 mg/kg	Hydrodolasetron, CYP2D6	Hydrodolasetron, CYP3A	5-HT3B
Palonosetron	Aloxi	40	0.75 mg	CYP2D6	CYP3A, CYP1A1/2	5-HT3A

Palonosetron

- Highly selective 2nd gen 5-HT3 Rc antagonist (>30 x Rc binding affinity vs 1st gen 5-HT3Ra's)
- Two stereogenic centers and may exist as four steroisomers
- Longer half-life (40 h)

Adverse effects of 5HT3 Rc Antagonists

Generally well-tolerated and side effects are infrequent

- Minor headache, flushing and constipations

- Higher doses used in chemotherapeutic induced nausea and vomiting (CINV) result in more serious side-effects – sedation, intestinal obstruction and incr ALT
  Ondansetron for PONV prophylaxis dose in adults is 4mg cf to CINV 32mg

- Ondansetron and dolasetron block sodium channels may result in **QRS widening**, and blocking potassium channels may lead to **QT prolongation** with resultant ventricular arrhythmias.

PSYCHIATRIC MEDICATION

Tricyclic Antidepressants (TCAs)

Used to treat depressive illness, nocturnal enuresis and as an adjunct in chronic pain therapy.
They competitively **block neuronal uptake of noradrenaline and 5-HT, Dopamine** thereby increasing the concentration of the transmitter in the synapse.

The antidepressant effects take up to 2 weeks to work.

They also block muscarinic, histaminergic and α-adrenoceptors, and have non-specific sedative effects.

Imipramine, Clomipramine and Trimipramine preferentially inhibit the reuptake of serotonin.

Monoamine Oxidase Inhibitors (MAOIs)

Used in the treatment of resistant depression, obsessive compulsive disorders, chronic pain syndromes and migraine.
MAO is present within presynaptic neurones and is responsible for the deamination of amine neurotransmitters. The main effect is rapid and sustained increase in 5HT, Dopamine and NA.

They have been classified as types A and B.

MAO-A preferentially deaminates 5-HT and catecholamines,
MAO-B preferentially deaminates tyramine and phenylethamine

The **original generation** inhibit MAO *irreversibly and non-selectively* (i.e.MAO-A and -B)
The **new generation** *selectively and reversibly inhibit* only MAO-A (RIMA).

Neither group is used as first line therapy because of the potential for serious side effects and hepatic toxicity.

Non Selective MAOIs
e.g. Isocarboxazid, Nialamide, Phenelzine, Hydracarbazine, Tranylcypromine

Tranylcypromine is potentially the most dangerous as it possesses stimulant activity.

Adverse effects
Sedation, blurred vision, orthostatic hypotension and hypertensive crises following tyramine rich foods (cheese, chocolate), indirectly acting sympathomimetics.(e.g. amphetamines, ephedrine)

Interaction with pethidine may precipitate cerebral irritability, hyperpyrexia and cardiovascular instability

Selective reversible MAOIs – RIMA (moclobemide)
Moclobemide , primarily used to treat depression and social anxiety, causes less potentiation of tyramine than the older generation MAOIs.

Linezolid is an antibiotic indicated for MRSA and vancomycin-resistant enetrococci. It is also a MAOI and as such has the typical range of cautions and contraindications.

Methylene blue (commonly used intraoperatively to delineate lymphovascular channels) is a highly potent, reversible MAOI.

MAOIs and general anaesthesia
Pethidine or any indirectly acting sympathomimetic amines (e.g. ephedrine) are contraindicated.

Indirect agents (e.g. dopamine) should be **avoided** because they are metabolized to adrenaline and noradrenaline. Under normal conditions, monoamine oxidase limits the intracellular concentration of these metabolites. When MAO is inhibited, Adrenaline and NA production at the cellular level is uncontrolled, possibly leading to an exaggerated hemodynamic response.

Direct-acting agents should be used for **CVS support** as they are metabolized by COMT and therefore not subject to the same degree of exaggerated response. Nevertheless, used with extreme caution as they may also precipitate exaggerated hypertension.

To diminish its effects, MAOI therapy should be withdrawn for 14–21 days before surgery (2 weeks before starting alternative therapy). This however, risks the patient suffering a relapse of their depression. However, the newer agents may control depression more effectively and reduce the chance of a serious peri-operative drug interaction.

Selective Serotonin Reuptake Inhibitors (SSRIs)

Globally, the most widely prescribed antidepressants e.g. Fluoxetine(Prozac ®),Citalopram(Cilift ®), Sertraline (Zoloft®).
There is selective inhibition of the neuronal re-uptake of 5-HT with minimal effect on dopamine and noradrenaline.

Although they are no more effective than standard antidepressants, they do not have their associated side-effect profile. (i.e. less sedative, fewer anticholinergic effects and appear less cardiotoxic in overdose although they are associated with gastrointestinal side-effects.)

5HT2 Antagonists: Antipsychotics

The main categories are **typical** (e.g. chlorpromazine and haloperidol) and **atypical** antipsychotics (e.g clozapine and risperidone). The distinction is unclear but rests on receptor action, incidence of extra-pyramidal side effects and efficacy.

5-HT2A and 5-HT2C receptor antagonists which cross the blood-brain barrier have antipsychotic properties and are called atypical neuroleptic or atypical antipsychotic agents

Clozapine is the reference drug of atypical antipsychotics.

In addition to its 5-HT2 antagonist effect, clozapine has several effects: antidopaminergic, antimuscarinic, antihistaminic and alpha-blocking. Its use is limited because of side effects such as agranulocytosis and potential to induce myocarditis.

The affinity for 5HT2 Rcs varies and is reflected here in their decreasing level of affinity; Clozapine, Resperidone, Flupenthixol, Chlorpromazine, Thioridazine and Haloperidol.

SEROTONIN SYNDROME (SS)

Is a potentially life- threatening, adverse drug reaction where there is increased serotonergic activity in CNS. It is seen with therapeutic medication use, inadvertent interactions between drugs, and intentional self-poisoning. In patients experiencing chronic pain there is a high incidence of depression and often on dual medication. These medications modulate serotonergic pathways compounding the risk of SS.

It is classically described as the **triad** of **Mental status changes, Autonomic hyperactivity, and Neuromuscular abnormalities (MAN).** However, it is a **spectrum of clinical findings** ranging from benign (barely perceptible) to lethal. Mild symptoms could be easily overlooked and inadvertent increase in dose of the culprit agent or addition of a proserotonergic drug may precipitate dramatic clinical deterioration.

Mental status changes can include anxiety, agitated delirium, restlessness, and disorientation Patients may startle easily. **Autonomic manifestations** can include diaphoresis, tachycardia, hyperthermia, hypertension, vomiting, and diarrhea. **Neuromuscular hyperactivity** can manifest as tremor, muscle rigidity, myoclonus, hyperreflexia, and bilateral Babinski sign.

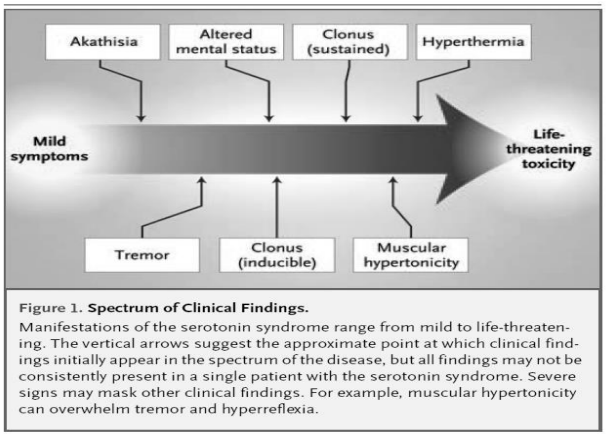


Figure 1. **Spectrum of Clinical Findings.**
Manifestations of the serotonin syndrome range from mild to life-threatening. The vertical arrows suggest the approximate point at which clinical findings initially appear in the spectrum of the disease, but all findings may not be consistently present in a single patient with the serotonin syndrome. Severe signs may mask other clinical findings. For example, muscular hypertonicity can overwhelm tremor and hyperreflexia.

N Engl J Med 2005;352:1112-20

Pathophysiology

Stimulation of the postsynaptic 5-HT1A/2A receptors in the central grey nuclei and medulla has been implicated in serotonin syndrome.

Noradrenergic CNS hyperactivity may play a critical role since the degree to which CNS concentrations are increased may correlate with clinical outcome.

Symptoms manifest predictably, with rising intrasynaptic serotonin concentration.
Intrasynaptic serotonin excess results from its increased synthesis, decreased reuptake or decreased metabolism.

Epidemiology and Incidence

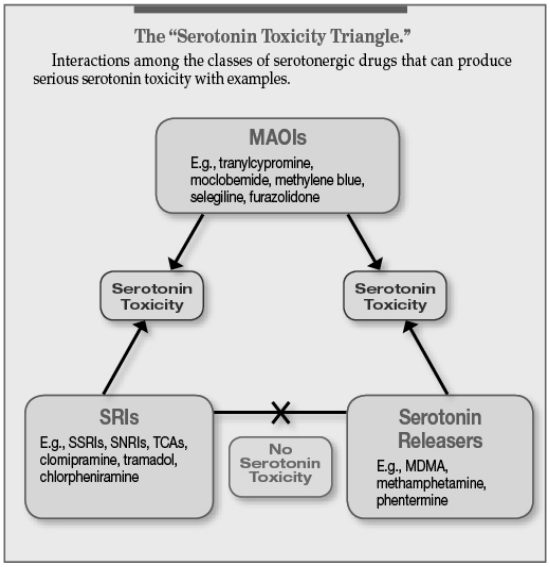
It is observed in all age groups, including newborns and the elderly. The increasing incidence of this condition is thought to mirror the increasing use of serotonergic agents in medical practice. However, the incidence may be under-reported for several reasons. Manifestations may be wrongly attributed to another cause, mild cases (such as tremor or with diarrhea or hypertension) may be viewed as inconsequential, or clinicians may just not suspect the condition.

There are numerous drug combinations associated with SS. The syndrome is classically associated with the simultaneous administration of two serotonergic agents, but it can occur after initiation of a single serotonergic drug or increasing the dose of a serotonergic drug in individuals who are particularly sensitive to serotonin.

SSRIs are perhaps the **most commonly implicated group** of medications associated with SS.

They may contribute to the development of SS up to several weeks after the drug has been discontinued. The half-life of fluoxetine is one week and that of its metabolite norfluoxetine is up to 2.5 weeks.

The **drug combination most commonly associated** with severe reactions is that of **MAOI and SSRIs**. Episodes of SS involving a **MAOI** may be more severe and more often lead to adverse outcomes, including death.



2015 Anaesthesia Patient Safety Foundation (5)

Serotonergic Opioids

Opioids remain a mainstay of pain management and are commonly used analgesics prescribed in conjunction with other serotonergic medications. Therefore, there exists a potential to influence the pharmacodynamics of serotonin.

Mechanisms for serotonergic action:

- weak serotonin reuptake inhibition and an
- increased release of intrasynaptic serotonin (through inhibition of GABA presynaptic inhibitory neuron on serotonin neurons)

Synthetic piperidine opioids - proserotonergic

Fentanyl
Augments serotonin release, and is a weak serotonin reuptake inhibitor (SRI)
Requires the additive action of other serotonergic agents or use in combination with SSRIs

Methadone
Greater tendency toward SRIn compared with other opiates.

Tramadol (racemic mixture)
RR enantiomer is a mu receptor agonist, while the SS enantiomer inhibits reuptake of norepinephrine and releases serotonin. Thus, resulting in increased intrasynaptic serotonin in combination with other serotonergic drugs

Dextromethorphan binds 5-HT2A receptors that may cause serotonin syndrome

Phenantherene morphine analogues – may increase the intrasynaptic serotonin levels either through increased release of neurotransmitter
e.g. oxycodone, hydromorphone, oxymorphone, and buprenorphine

Table 2. Serotonergic Opioids Association in Serotonin Syndrome

Opioid	Coadministered Agent
Tramadol	Paroxetine, fluoxetine
Oxycodone	Escitalopram, sertraline, fluvoxamine, citalopram
Fentanyl	Escitalopram, trazadone, paroxetine, venlafaxine, citalopram, sertraline, granisetron, dihydroergotamine
Methadone	Venlafaxine, sertraline, fluoxetine, linezolid
Dextromethorphan	Bupropion, methadone, citalopram, sertraline, escitalopram, chlorpheniramine, tramadol
Meperidine	Citalopram, fluoxetine, linezolid, moclobemide, venlafaxine
Codeine	Sertraline
Buprenorphine	Doxepin, amitriptyline

Assessment of Patient

History
A detailed **description** of prescription **drugs**, over-the- counter medications (cough syrup ingredient dextromethorphan), illicit substances (ecstasy), and dietary supplements (weight loss drugs fenfluramine and sibutramine), as well as any changes in dosing and formulation (e.g. sustained release) must be recorded.

An account of **symptoms,** their evolution and rate of change should be sought. The majority of cases of SS present within 24 hours(75%), and most within six hours (60%), of a change in dose or initiation of a drug. Many cases typically resolve within 24 hours after the initiation of therapy and the discontinuation of serotonergic drugs. Symptoms may persist where patients are taking drugs with long elimination half-lives, active metabolites, or a protracted duration of action.

Physical examination
Refer to MAN acronym above

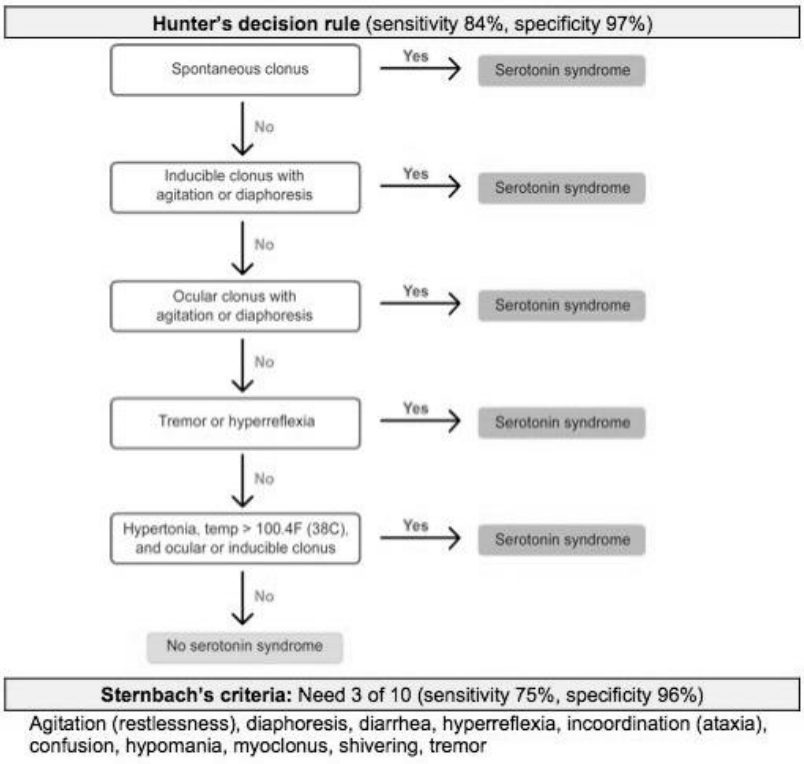
Laboratory Evaluation
Are of very little use in the diagnosis as serum serotonin concentrations do not correlate with clinical findings.

Other findings are nonspecific - elevated white blood cell count, elevated CK, and decreased serum bicarbonate concentration.

In severe disease, there may be development of profound complications- including metabolic acidosis, DIC, rhabdomyolysis, renal failure, myoglobinuria, and ARDS.

Diagnosis and Diagnostic Criteria

Hunter Toxicity Criteria is most accurate (84 percent sensitive and 97 percent specific)



Ables AZ, Nagubilli R. *Am Fam Physician*. 2010 May 1;81(9):1139-42.

Differential Diagnosis

Includes neuroleptic malignant syndrome (NMS), anticholinergic toxicity, malignant hyperthermia, intoxication from sympathomimetic agents, sedative-hypnotic withdrawal, meningitis, and encephalitis. Each of these can be differentiated from SS on clinical grounds. Neuroleptic syndrome can be precipitated by use of antipsychotic medications and may present with similar symptoms.

Serotonin syndrome and neuroleptic malignant syndrome: Distinguishing features

	Serotonin syndrome (SS)	Neuroleptic malignant syndrome (NMS)
Onset	Within 24 hours	Days to weeks
Neuromuscular findings	Hyperreactivity (tremor, clonus, reflexes)	Bradyreflexia, severe muscular rigidity
Causative agents	Serotonin agonist	Dopamine antagonist
Treatment agents	Benzodiazepine, cyproheptadine	Bromocriptine
Resolution	Within 24 hours	Days to weeks

Table 4. Differential Diagnosis of Serotonin Syndrome

Clinical Conditions	Distinguishing Features
Serotonin syndrome (serotonin excess)	Cognitive: anxiety, agitation, confusion, hypomania, visual hallucinations, restlessness, disorientation, and coma. Autonomic: hyper/hypotension, tachycardia, tachypnea, diarrhea, mydriasis, diaphoresis, and hyperthermia. Neuromuscular: muscle rigidity, tremors, nystagmus, myoclonus, ocular clonus, hyperreflexia, ataxia, and trismus
Neuroleptic malignant syndrome (dopamine antagonism)	Extrapyramidal symptoms, "lead pipe" rigidity, gradual onset, bradykinesia, absence of gastrointestinal hyperactivity, myoclonus, and hyperreflexia
Anticholinergic syndrome (cholinergic antagonism)	History of anticholinergic agent use (such as tricyclic antidepressants), widened pulse pressure, dry skin and mucus membranes, normal reflexes, absence of myoclonus, and gastrointestinal hyperactivity
Malignant hyperthermia	History of halogenated anesthetic and depolarizing muscle relaxant exposure, hyporeflexia, and absence of myoclonus
Opioid toxicity	History of opioid exposure, miosis, hypotension, hypothermia, bradycardia, hypopnea, and hyporeflexia
Opioid withdrawal	History of sudden opioid discontinuation or intake of opioid antagonist agents, piloerection, joint pains, "flu-like" symptoms, absence of hyperreflexia, and myoclonus

Management: key principles

- Discontinuation of all serotonergic agents
- Supportive care aimed at normalization of vital signs
- Sedation with benzodiazepines
- Administration of serotonin antagonists
- Assessment of the need to resume use of causative serotonergic agents after resolution of symptoms

Application of these principles varies with the severity of illness. Common management pitfalls include failure to recognize serotonin syndrome, misdiagnosis and failure to appreciate SS's potentially rapid rate of progression

Mild cases
Discontinuation of inciting medications, supportive care, and sedation with benzodiazepines are generally sufficient.

> Supportive care and sedation is the mainstay of therapy
> oxygen and intravenous fluids, continuous cardiac monitoring, and correction of vital signs
>
> Chemical restraint is preferred over physical restraint for agitated patients; physical restraints may cause isometric muscle contractions leading to profound lactic acidosis and hyperthermia. Sedation with benzodiazepines is important for controlling agitation as well as blunting the hyperadrenergic component of the syndrome. Improved survival in animal models has been demonstrated.
>
> Butyrophenones (eg, droperidol and haloperidol) should be avoided; these drugs have anticholinergic properties that inhibit sweating and dissipation of body heat.

Moderately ill patients
More aggressive treatment of autonomic instability and possibly treatment with a serotonin antagonist (see antidote)

Critically ill
These patients are hyperthermic (>41.1°C) and often require neuromuscular paralysis and tracheal intubation. Suxamethonium should be avoided because of its detrimental effects in the context of rhabdomyolysis.

There is no role for antipyretic agents, such as acetaminophen since the temperature elevation is not due to an alteration in the hypothalamic temperature set point, but rather an increase in muscular activity

Autonomic instability
The challenge in severely intoxicated patients is that they often exhibit large and rapid changes in blood pressure and heart rate.
In patients with severe hypertension and tachycardia, a strategy of **titrating short-acting cardiovascular agents**, such as esmolol or nitroprusside to maintain autonomic stability, may be successful.

Hypotension from MAOIs in patients with SS should be treated with low doses of direct-acting sympathomimetic amines, such as phenylephrine, epinephrine, or norepinephrine
Indirect agents (e.g. dopamine) should be avoided (see explanation above in PSYCH MEDS MAOIs)

Antidote
If benzodiazepines and supportive care fail to improve agitation and correct vital signs

Cyproheptadine

- Histamine-1 receptor antagonist weak anticholinergic activity
- Non-specific 5-HT1A and 5-HT2A antagonistic properties - may produce transient hypotension due to the reversal of serotonin-mediated increases in vascular tone. Such hypotension usually responds to intravenous fluids.

- Only available in an oral form, but it may be crushed to administer through a nasogastric tube (per oral form in 2 mg incremental doses for a maximum of 12–32 mg in 24 h).

Definitive evidence of cyproheptadine's effectiveness is lacking

Other antidotes — Antipsychotic agents with 5-HT2A antagonist activity, such as olanzapine and chlorpromazine, have been considered for antidotal treatment, but their efficacy is unproven and NOT recommended. Chlorpromazine can cause orthostatic hypotension and increase hyperthermia.

Treatment with propranolol, bromocriptine, or dantrolene is not recommended.
Propranolol has a long duration of action, may cause prolonged hypotension, and can mask tachycardia that can be used to monitor the effectiveness of treatment. Bromocriptine, a serotonin agonist, may exacerbate serotonin syndrome. Dantrolene has no effect on survival in animal models.

Fig. 2. Spectrum of serotonin syndrome and their management strategies.

CONCLUSION

Anaesthetists need to be aware of the increased risk of serotonin syndrome in the perioperative setting, where high dose opioids are often administered. Similarly, practitioners in chronic pain management, where use of opioids in conjunction with other serotonergic agents is the norm, must also have heightened sense of awareness.

Prognosis is generally favorable, as long as the entity is recognized and complications are treated appropriately.

References

1. Allain P. Serotonin, serotoninomimetics, antagonists. Retrieved on 25 February 2017 from http://www.pharmacorama.com/en/Sections/Serotonin_2_1.php
2. Boyer EW. The Serotonin Syndrome. N Engl J Med 2005; 352(11): 1112-1120.
3. Boyer EW. Serotonin syndrome (serotonin toxicity).UpToDate. Retrieved 04 April 2017 from https://www.uptodate.com/contents/serotonin-syndrome
4. Chinniah S et al. Serotonin and anaesthesia. CEACCP 2008; 8: 43-45
5. Gyermek L. Pharmacology of Serotonin Related to Anesthesia. J of Clin Anesth 1996; 8: 402-425
6. Locke A. Methylene Blue and the Risk of Serotonin Toxicity. June 2015. Retrieved on 14 April 2017 from http://www.apsf.org/newsletters/html/2015/June/02blue.htm
7. Mulder R. Serotonin A mediator of our times Part I Anaesthetic Refresher Course 2010, 14:1-12
8. Peck T. Pharmacology for Anaesthesia and Intensive Care. 3rd Edition, 2008. Cambridge University Press.
9. Rastogi R. Case Scenario: Opioid Association with Serotonin Syndrome. Anesthesiology 2011; 115:1291– 8
10. Walsh J. Serotonin Syndrome. Anaesthesia Tutorial of the Week 166, 2010. Wfsahq.org. Retrieved on 04 April from http://www.frca.co.uk/Documents/166%20serotonin%20syndrome.pdf

Basic Definitions in Statistics

Dr S.A.M. Heijke

UCT Dept of Anaesthesia & Perioperative Medicine

The fundamental idea of statistical analysis of medical studies is that we make observations on a sample of subjects and then draw inferences about the population from which the sample was drawn.

Measures of position

The mean, median and mode are measures of **location or** central tendency i.e. give an idea of *where* the middle of the information is.

MEAN:

The mean is the arithmetic average.
Add all the numbers and divide by the number of numbers.

$$\textbf{Average} = \frac{\textbf{Sum of observations}}{\textbf{Number of observations}}$$

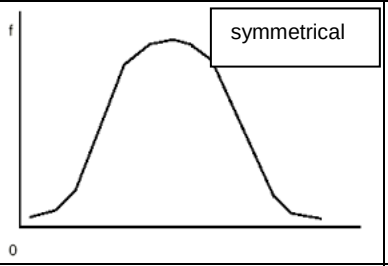
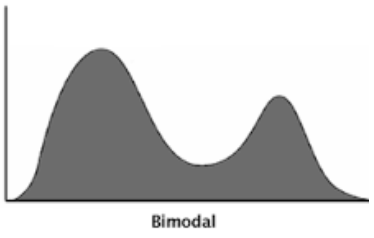
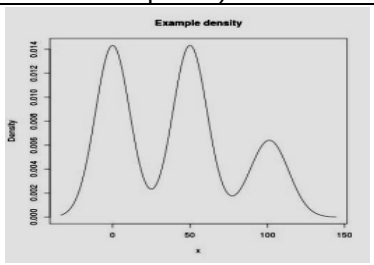
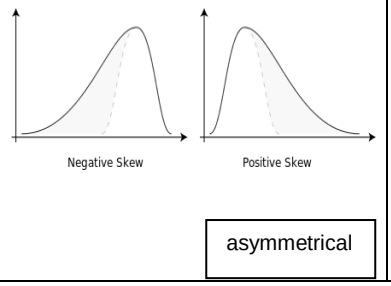
MEDIAN:

The value or quantity lying at the midpoint of a frequency distribution, i.e.

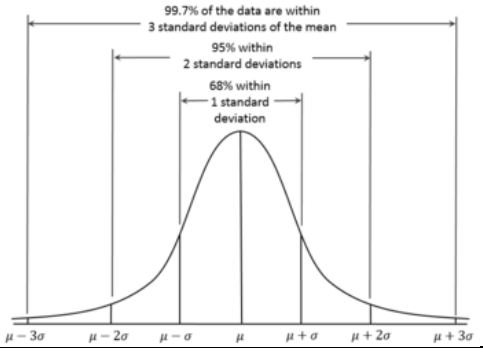
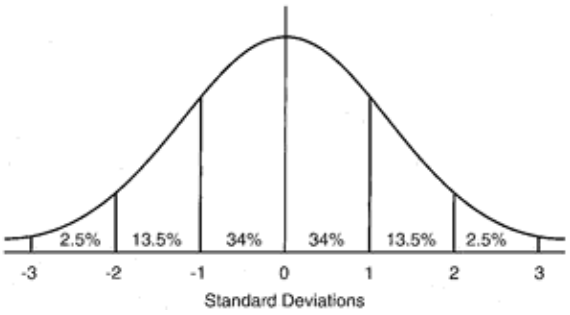
= the value of the middle observation if the number of observations is odd or
= the value of the average of the middle two observations if the number of observations is even

MODE:

The value or values that occur most frequently in the set of observations. i.e. results may be unimodal, bimodal or multimodal.

Unimodal (one peak)	Bimodal (two peaks)	Multimodal (two or more peaks)
		
		

A normal distribution: has the same mean, mode and median. The Standard Normal distribution has a normal distribution with a mean of 0.

Normal distribution	Standard Normal distribution
	

Measures of spread

The range, variance, standard deviation and the standard error of the mean give information about the degree of scatter or spread of the data.

RANGE:

The size of the interval that contains all the data in descriptive statistics or arithmetically the highest value minus the lowest value.

VARIANCE:

The variance measures how far the data points are from the mean. If you simply added them up the sum would be Zero because some lie on the positive side and some on the negative side so we can either take the absolute values of the distance from the mean, but this is difficult to work with so instead we can square the values of this difference and then add them together. To get the variance we then divide this sum by the number of observations n minus 1, i.e. n-1. This is because we are calculating the variance of a sample. If we were calculating the variance of the population, i.e. the entire set of data we would divide by N and we would use the Greek letter sigma or σ.

$$S^2 = \frac{\Sigma(X_i - \bar{X})^2}{n-1}$$

S - Variance

$\sum$ - Sum of

X_i - Data value

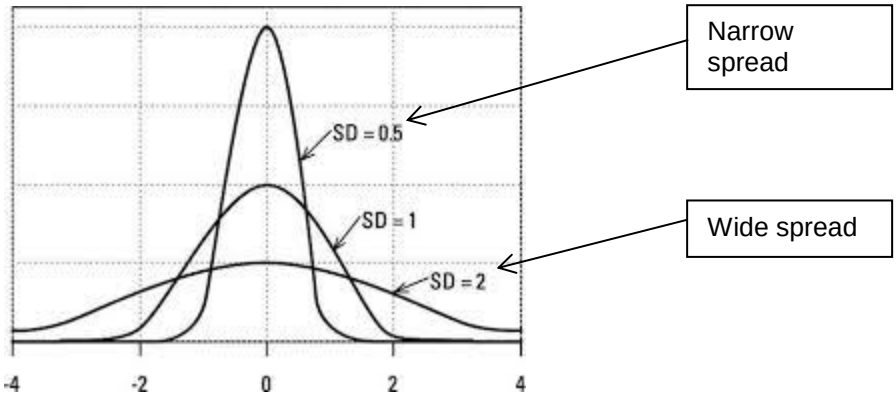
$\bar{X}$ - The mean

n - The number of observations

STANDARD DEVIATION (SD):

The standard deviation is the square root of the variance

$$S = \sqrt{\frac{\Sigma(X_i - \bar{X})^2}{n-1}}$$



STANDARD ERROR OF THE MEAN (SEM):

The standard error of the mean estimates how close the sample mean is to the population mean. It measures the variability of the sample means. The lower the value the more precisely the sample means reflect the true population mean. The lower the variance or the higher the number of observations the smaller this error will be.

$$SEM = \frac{S}{\sqrt{n}}$$

Biostatistics terms

PREVALENCE:

The percentage or proportion of the population that has a specific disease at a specific time, i.e. how common or widespread is the disease or condition (new and pre-existing cases).

$$Prevalence\ Rate = \frac{\begin{array}{c}Total\ number\ of\ new\ and\\ pre-existing\ cases\ of\ a\ disease\\ during\ a\ given\ time\ period\end{array}}{\begin{array}{c}Total\ population\ during\ the\\ same\ time\ period\end{array}} * 10^n$$

INCIDENCE:

The incidence indicates what the risks are of getting a disease (new cases).

$$Incidence\ Rate = \frac{\begin{array}{c}Total\ number\ of\ new\ cases\ of\ a\ disease\\ during\ a\ given\ time\ period\end{array}}{\begin{array}{c}Total\ population\ at\ risk\ during\ the\\ same\ time\ period\end{array}} * 100^n$$

E.g. A disease that is chronic may have a low incidence but high prevalence and a disease that has a short duration may have a high incidence but a low prevalence.

The Truth Table

		The "Truth"	
		Yes	No
Test Result	Yes	(A) True +	(B) False +
	No	(C) False –	(D) True –

Sensitivity: The ability of a test to correctly identify persons with the disease, i.e. detect the true positive rate.

$$\frac{tp}{tp + fn}$$

Specificity: The ability of a test to correctly identify persons without the disease, i.e. detect the true negative rate.

$$\frac{tn}{fp + tn}$$

True negative: Does not have the disease and tests negative for the disease.

False positive: Does not have the disease but tests positive for the disease.

		True class: Positive	True class: Negative	Measures
Predicted class	Positive	True positive *TP*	False positive *FP*	Positive predictive value (PPV) $\frac{TP}{TP+FP}$
	Negative	False negative *FN*	True negative *TN*	Negative predictive value (NPV) $\frac{TN}{FN+TN}$
Measures		Sensitivity $\frac{TP}{TP+FN}$	Specificity $\frac{TN}{FP+TN}$	Accuracy $\frac{TP+TN}{TP+FP+FN+TN}$

POSITIVE PREDICTIVE VALUES (PPV):

Positive predictive value is the probability that subjects with a positive screening test truly have the disease.

$$\frac{tp}{tp + fp}$$

NEGATIVE PREDICTIVE VALUES (NPV):

Negative predictive value is the probability that subjects with a negative screening test truly don't have the disease.

$$\frac{tn}{fn + tn}$$

RELATIVE RISK AND ODDS RATIO:

The relative risk (RR) is the ratio of the risk of an outcome in an exposed or treated group compared to a control group, e.g. how likely are you to get lung cancer if you smoke compared to if you do not. If there is no difference the risk will be one, if the risk is lower the ratio will be <1 and if it is higher the value will be >1.

The odds ratio (OR) is a measure of the association between exposure and an outcome. It represents the chance or **odds** that an outcome will occur given a particular exposure, compared to the **odds** of the outcome occurring in the absence of that exposure.

Odds Ratio (OR)

Contingency (or 2 x 2) Table

	Cases	Controls	Total
Exposed	a	b	a+b
Unexposed	c	d	c+d
Total	a+c	b+d	a+b+c+d

$$OR = (a/c) / (b/d) = (a*d) / (b*c)$$

Outcome

Predictor

	Yes	No
Yes	A	B
No	C	D

$$OR = \frac{A*D}{B*C}$$

If the prevalence of the disease is low then the odds ration approaches the relative risk

CONFIDENCE INTERVAL:

A confidence interval is a **range** of values within which we are fairly sure our true value lies. Usually we use 95%. This means a 95% "chance" that our value will be within this range and 5% that it will lie outside. If there are values above and below the range (two-tailed), then there will be 2.5% chance that the value will lie outside above the range and 2.5% chance that the value will be below the range. If the values are to one side only (one-tailed) then there will be a five percent chance that our result will fall there.

The mathematical calculation of this confidence interval depends on what we are measuring. For the mean we use a Z value (from tables)

Mean $\bar{x} \pm Z\frac{s}{\sqrt{(n)}}$ Margin of error

The size of the confidence interval depends on the variation within the population and the size of the sample

The size of the confidence interval depends on the variation within the population and the size of the sample

P-VALUE:

Probability is the likelihood of an event occurring by chance

"The **P value**, or calculated probability, is the probability of finding the observed, or more extreme, results when the null hypothesis (H_0) of a study question is true – the **definition** of 'extreme' depends on how the hypothesis is being tested."

In every experiment the researcher compares two or more groups. The null hypothesis (H_0) is the possibility that there is **no** difference between the groups being tested. The alternative hypothesis (H_1) is that any differences are real. Random sampling error may however show a difference between groups even if there really is not.

A small **p-value** (typically ≤ 0.05) indicates strong evidence against the null hypothesis, so you reject the null hypothesis. A large **p-value** (> 0.05) indicates weak evidence against the null hypothesis, so you fail to reject the null hypothesis.

- High P values: your data are likely with a true null.
- Low P values: your data are unlikely with a true null.

Setting up experiment steps:

1) State Hypothesis – Null (H_0) versus alternative (H_1)
2) Decide Significance level or α level – commonly 0.05
3) Collect sample
4) Calculate your p value (using tables or computer program)
5) Decide: if your value is less than p, reject the null hypothesis
   if your value is more than p, then accept the null hypothesis

STATISTICAL ERRORS:

	Reject H_0	Accept H_0
H_0 is true	Type 1 error α error	Correct
H_0 is false	Correct	Type 2 or β error

References

1. Numerous YouTube clips that are simplified for easier understanding though professional statisticians may be unhappy. e.g. P VALUE: https://www.youtube.com/watch?v=eyknGvncKLw
2. The mathsisfun site is easy to use and understand and has a great section on data mathematics https://www.mathsisfun.com/data/index.html

Statistical Graph Interpretation

Prof. Bruce Biccard

UCT Dept of Anaesthesia & Perioperative Medicine
bruce.biccard@uct.ac.za

Introduction

In this tutorial, we will cover the common statistical graphs one encounters in clinical medicine. We will discuss the indications and interpretation of the graphs, as important considerations when appraising the validity (and limitations) of the graph for the presentation of the data. This tutorial is built predominantly around the following works.[1-4]

Graphs used to describe the characteristics of data

1. Histograms

Indication: To show the frequency and the shape of continuous data. Determining whether data is normally distributed (or can be transformed to normality) allows for the use of parametric tests in data analysis.

Interpretation: A histogram allows for a visual inspection of the data distribution. It is possible therefore to determine whether it approaches a bell (or Gaussian) distribution. It shows where there are gaps (no data points), outliers, and whether the data is skewed. The skewness is named by the longer tail i.e. positive skew, has a long right tail.

Limitations: Normality of data cannot be determined by visual inspection of a histogram alone. Other formal statistical tests are needed to describe normality. The reasons for a deviation from normality may be statistically described by the skewness (positive or negative skew) and kurtosis (flat or pointed distribution) of the data.

Data organisation: The data categories are on the X-axis, and the frequency of the data in each category is plotted on the Y-axis.

When you are most likely to use a histogram: Histograms are infrequently presented in peer-reviewed papers. It is most likely that you will use a histogram as part of your exploratory work on your data to determine if it is normally distributed in preparation for the choice of an appropriated statistical test for the analysis your MMed.

Classic example in the anaesthesia literature: One of the original papers describing the anaerobic threshold (AT) in preoperative surgical patients by Paul Older and colleagues presented the following histogram of the distribution of the AT in their surgical patients.[5] In published data it is important to calculate the 95% confidence interval (CI) from the standard deviation (SD) (95% CI=mean $\pm$ 2 SD), to determine if the data presented are actually plausible. In this example, the average oxygen consumption for all patients at AT was 12.4ml/kg/min with a SD=2.7,[5] therefore the 95% CI for the AT was 6 to 17.8ml/kg/min. This is entirely plausible. The data looks normally distributed (with a possible small negative skew), yet we would need to confirm normality with further statistical analysis.

Other complementary graphs of data normality: The Q-Q plot. Here, the observed value (X-axis) is plotted against the expected value (Y-axis). Deviations from a straight line suggest deviation from normality of distribution.

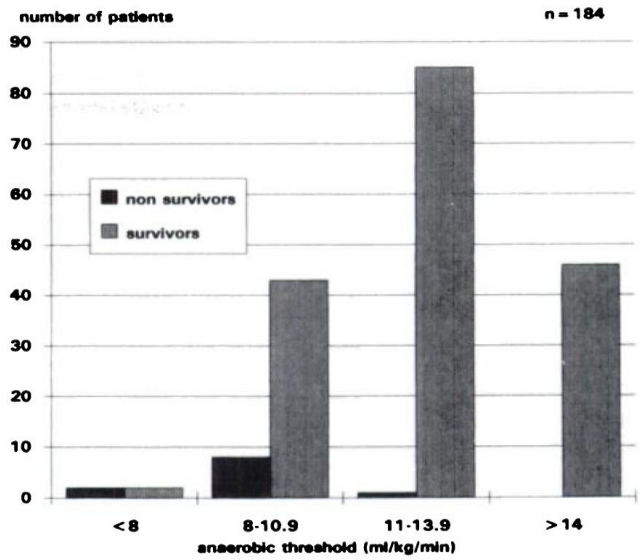


FIGURE 3. Comparison of cardiovascular mortality with anaerobic threshold.

Figure 1. Histogram presented by Paul Older and Colleagues describing the anaerobic threshold in surgical patients[5]

2. Box and whisker plots

Indication: To graphically present the median and interquartile range (IQR) in non-normally distributed data.

Data organisation: The median is presented as the horizontal line in the box, and the length of the box represents the IQR. Extreme values are usually shown when they are more than three box lengths from the upper or lower box end.

When are you most likely to use a box and whisker plot: Presentation of data which is non-normally distributed. This is common with patient biochemical data. The typical presentation is a positive skew due to small proportion of patients having, for example a high serum creatinine, troponin or B-type natriuretic peptide level compared to the population. It is possible to transform non-normally distributed data to a normal distribution in order to use parametric statistical tests. The most common transformation is a log transformation for positively skewed data. This is commonly used with data such as B-type natriuretic peptide levels which are positively skewed.

Classic example in the anaesthesia literature: Cuthbertson and colleagues presented the association between preoperative B-type natriuretic peptides (BNP) and postoperative cardiac events.[6]

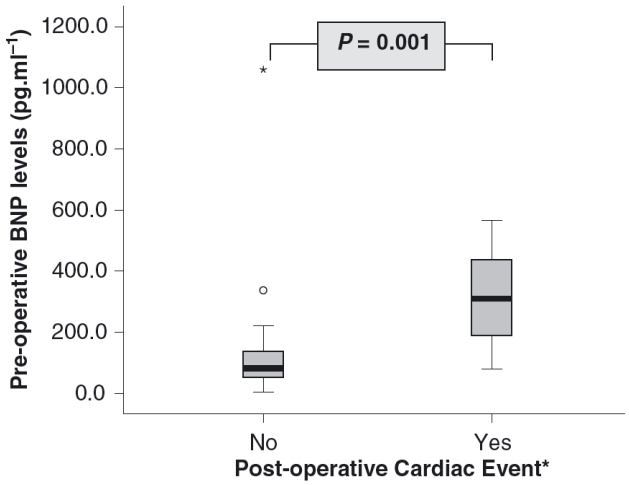


Figure 1 A comparison of pre-operative BNP levels in patients who experienced a postoperative cardiac event with those who did not. Central lines represent medians, boxes 25th and 75th centiles and whiskers represent ranges.

Figure 2. The association between preoperative B-type natriuretic peptides (which are non-normally distributed and positively skewed) and postoperative cardiac events.[6]

Critical appraisal of data normality:

1. Has the normality of the continuous data been determined both visually and statistically?
2. Has the centre and distribution been described appropriately?
3. If the data is claimed to be normally distributed, does the mean±2SD describe a plausible 95% confidence interval?

Graphs used to describe the relationship between two variables

1. Scatter plots or scattergrams

Indication: To understand the nature of the relationship between two continuous variables. This describes how closely two variables are relate, and the amount of variability in one measurement that can be explained by another measurement.

Interpretation: To express the numerical value of the correlation, we calculate the correlation coefficient. The correlation coefficient quantitively expresses both the magnitude (from 0 to 1) and the direction of the correlation (positive i.e. both increase in value together, or negative i.e. one variable deceases, while the other increases). The proportion of variance that can be attributed to one variable, is dependent of the 'coefficient of determination', which is the square of the correlation coefficient e.g. r=0.7, therefore coefficient of determination (r^2) = 0.49 i.e. 49% of the variation can be explained by the relationship between the two variables, and 51% is due to other factors.

Limitations: Firstly, correlation does not represent causation. Causality is dependent of three criteria; i) the causative occurrence needs to precede the effect, ii) covariation of cause and effect i.e. if the cause occurs, then the effect should occur, and iii) elimination of rival causal explanations i.e. if you remove the cause, then the effect should not occur. This is why we use randomised controlled trials to determine causality. Secondly, a lack of correlation using the above tests does not necessary reflect no correlation, as these tests only reflect linear correlation. The two variables may be correlated in another way i.e. some curved relationship.

Data organisation: This is dependent on the type of data. For linear correlations only, the commonly used tests are; i) Pearson's (r) correlation coefficient, which plots two continuous variables (interval or ratio scale) which are normally distributed, ii) Spearman's (rho) correlation coefficient, which plots either two ordinal variables, or one continuous, and one ranked variable, and iii) Kendall's correlation coefficient, which plots two categorical variables.

When are you most likely to use a correlation scattergram: For testing the correlation between tow variables e.g. between two different blood tests measuring the same serum blood marker.
Classic example in anaesthesia literature: In order to determine the association between perioperative volume expansion, and a >15% increase in cardiac output, Yannick le Manach and colleagues look at the association between various haemodynamic variables (change in SBP, DBP, MAP and pulse pressure (PP) variation).[7] The only significant association between volume expansion and an increase in cardiac output of >15% was a > 3% decrease in the PPV i.e PPV after volume expansion (VE) − PPV before VE. The correlation between the change in the PPV and the CO is shown in the figure. The association is negative i.e. the PPV falls, for an increase in the CO. The change in PPV can explain 36% of the change in the CO in this study.

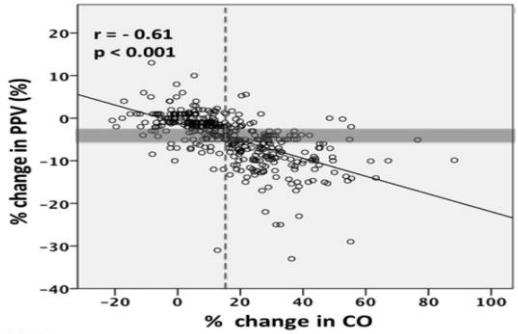


Fig. 2. Relationships between volume expansion–induced changes in PPV and CO. Percent change in PPV = percent changes in pulse pressure variation (defined as pulse pressure variation after volume expansion − pulse pressure variation before volume expansion). Percent change in CO = percent changes in cardiac output induced by volume expansion. The *dashed vertical line* shows the 15% threshold increase in CO defining fluid responsiveness.

Figure 3. The correlation between pulse pressure variation and cardiac output[7]

2. Altman-Bland Plot

Indication: To quantify the agreement between two readings.

Interpretation: Incorrect methods of agreement between two readings are to; i) compare the means (and find no significant difference), or ii) the correlation coefficient (it is a measure of association, not a measure of agreement).

Limitations: This analysis cannot be used to predict one measurement from another. Prediction of a measurement is not adequately addressed by the Altman-Bland method of comparing measurements.
Data organisation: Plot the difference between the methods (A-B) on the Y-axis, against the average of the methods, (A+B)/2, on the X-axis. The mean of the difference (A-B) is the relative bias, and the SD is the estimate of the error.

When are you most likely to use an Altman-Bland plot: Comparing a measurement by a new device/ monitor, against the 'gold standard'.

Classic example in the anaesthesia literature: Suehiro and colleagues comparing 3D transoesophageal echocardiography (TOE) and pulmonary artery (PA) catheter estimation of stroke volume in cardiac surgical patients.[8] They found that the 3D TOE had a negative bias of 1.2ml for the stroke volume with a 95% CI of 11.9ml to -14.3ml).

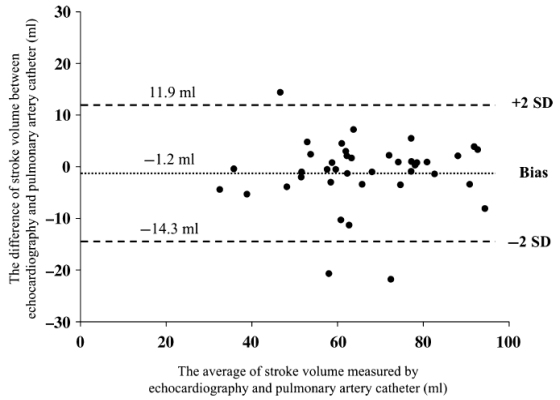


Figure 3 Bland–Altman analysis to investigate the reliability of 3D-transoesophageal echocardiography compared with the pulmonary artery catheter in measuring stroke volume.

Figure 4. Altman-Bland plot of the bias and estimated error of using 3D TOE to determine stroke volume in cardiac surgical patients

Critical appraisal of scatter plots:

1. No matter how small the p-value is, it is the r that is important
2. R is not on a linear scale, but r^2 is
3. A significant p-value, or a high r, do not imply causation.

Receiving operating characteristic (ROC) curves

Indication: To determine a cut-off point in continuously distributed data, that predicts the presence of an outcome. This cut-off point can be either; i) a screening cut point, ii) a diagnostic cut point, or iii) the optimal cut point.

Interpretation: ROC curves are useful for determining a screening cut point i.e. 95% sensitive (curve closest to the top of the Y-axis, where true positives are maximised), or a diagnostic cut point i.e. 95% specific (curve closest to the left of the X-axis, where the rate of false positives is low, and thus true negatives are high).

Limitations: Studies that only present the optimal cut, over-estimate the utility of a test.[9] This is because the data of a single study has been dichotomised to maximise sensitivity and specificity i.e. to determine the point on the ROC curve which is closest to the ideal point i.e. shortest distance to the top left corner of the plot i.e. Distance $=(1\text{-sensitivity})^2 + (1\text{-specificity})^2$

Data organisation: A plot of sensitivity i.e. true positives (on the Y-axis) against 1-specificity i.e. false

positives (on the X-axis). A move upwards shows a true positive result, and a move to the right shows a false positive result. Therefore, the ideal test will be at the top left hand corner of the plot i.e. 100% sensitive and 100% specific. The area under the diagonal line is 0.5 i.e. there is no difference from chance. The success of the test is reported as the area under the curve (AUC).

When are you most likely to use a ROC curve: To determine an appropriate cut point for a test e.g. at what STOP-BANG score should you consider postoperative apnoea a clinical problem.

Classic example in the anaesthesia literature: Choi and colleagues presented the improved discrimination shown by adding preoperative CRP and preoperative BNP to the Revised Cardiac Risk Index (RCRI) in predicting major adverse cardiac events following surgery.[10]

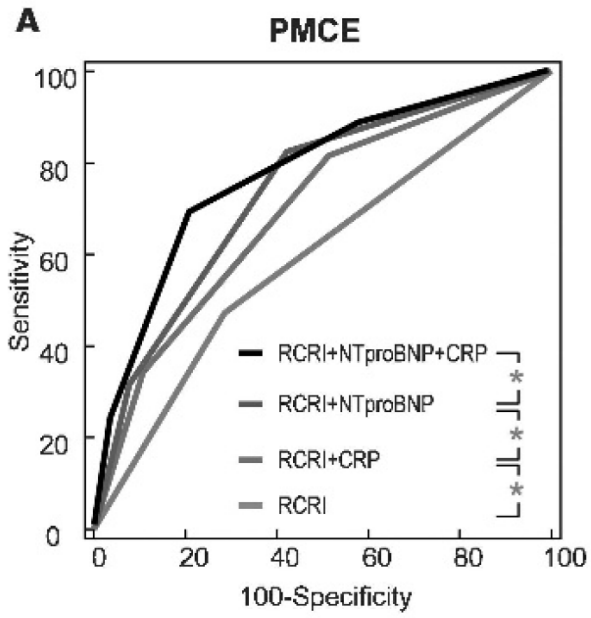


Figure 5. The improvement in discrimination when adding the preoperative CRP or BNP to the Revised Cardiac Risk Index for predicting postoperative major adverse cardiac events.

Critical appraisal of ROC curves:

1. Was the gold standard used to classify the diagnosis?
2. Were the results of the test withheld from the people who classified patients as having the disease (and vice versa)?
3. Are there enough test positive and test negative patients to accurately determine sensitivity and specificity?
4. Are there confidence intervals reported?

Survival plots (or Kaplan-Meier curves)

Indication: To present the time to an outcome in (usually) two different groups (or cohorts). The time to the primary outcome in randomised controlled trials are often presented in survival plots, which are known as Kaplan-Meier curves.

Interpretation: The utility of a survival plot is that it can indicate the time period at which a patient is most likely to be at risk of the outcome i.e. the steepest part of the curve.

Limitations: It is limited by; i) the duration of the follow up i.e. events that occur after the time of completion of follow up are not recorded, or ii) loss to follow up, where patients are censored as per their known last status (which is usually no reported outcome event).

Data organisation: The cumulative outcome is reported on the Y-axis, and the time of the event is

reported on the X-axis. Y-axis data is presented either descending i.e. number of individuals free of the outcome, or ascending i.e. number of individuals who have experienced an outcome event. Each step represents an additional outcome event.

When are you most likely to use a survival plot: To report the time of specific outcomes in two patient cohorts.

Classic example in the anaesthesia literature: The POISE[11] Trial which showed the different beneficial and harmful outcomes associated with perioperative acute beta-blockade, with decreased myocardial infarction and increased stroke and mortality within 30-days of surgery.

Critical appraisal of survival plots:

1. Is the follow up period long enough to identify the outcome?
2. Is there a significant number of patients lost to follow up during the follow up time period?
3. Has the scaling of the Y-axis been manipulated to increase the visual discrimination between groups?

Figure 6. POISE trial outcomes within 30 days of surgery following randomisation to beta-blockade or placebo.[11]

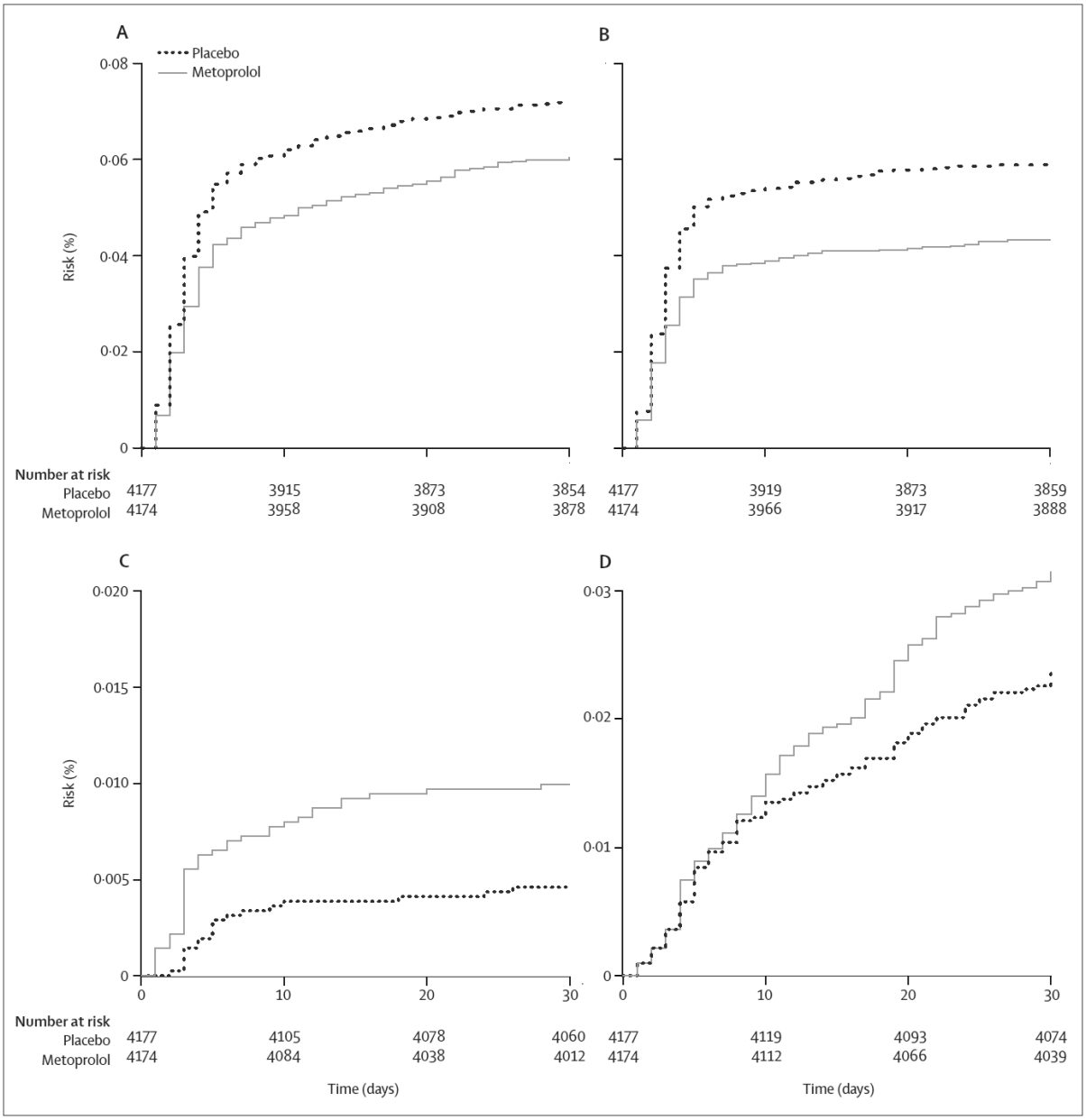


Figure 2: Kaplan-Meier estimates of the primary outcome (A), myocardial infarction (B), stroke (C), and death (D)

Forest plots

Indication: To present the summary data of a meta-analysis.

Interpretation: It summarises the current, aggregated data. If the systematic review was exhaustive, the meta-analysis provides the best evidence of the current data. This information, allows for determination of appropriate, future research projects within the field.

Limitations: Sometimes, there are too few data to allow for a meta-analysis. In this case the data can only be described in a systematic review.

Data organisation: The data for two groups (an intervention group and a control group) are presented. Data can be reported as binary outcomes e.g. alive or dead, or continuous data e.g. change in forced expiratory volume. Continuous data is usually presented as the weighted mean difference (WMD) between groups. The weighted mean difference can be converted back to the standard units of measurement in order to understand the clinical impact of an intervention. The Y-axis lists the studies/ trials included in the meta-analysis. The X-axis shows the line of unity, and the association with the outcome, either as the relative risk (RR), odds ratio (OR), or weighted mean difference (WMD).

When are you most likely to use a forest plot: In a systematic review where there is enough similar data to allow aggregating data from more than one study.

Classic example in the anaesthesia literature: The meta-analysis of perioperative beta-blockade, which conducted subgroup analyses of the secure and insecure studies (following the fraudulent findings against Don Poldermans).[12] The data showed high heterogeneity and treatment effect in the insecure trials, when compared to the secure trials.

Figure 7. A meta-analysis of the effect of acute perioperative beta-blockade on myocardial infarction in secure and insecure clinical trials.[12]

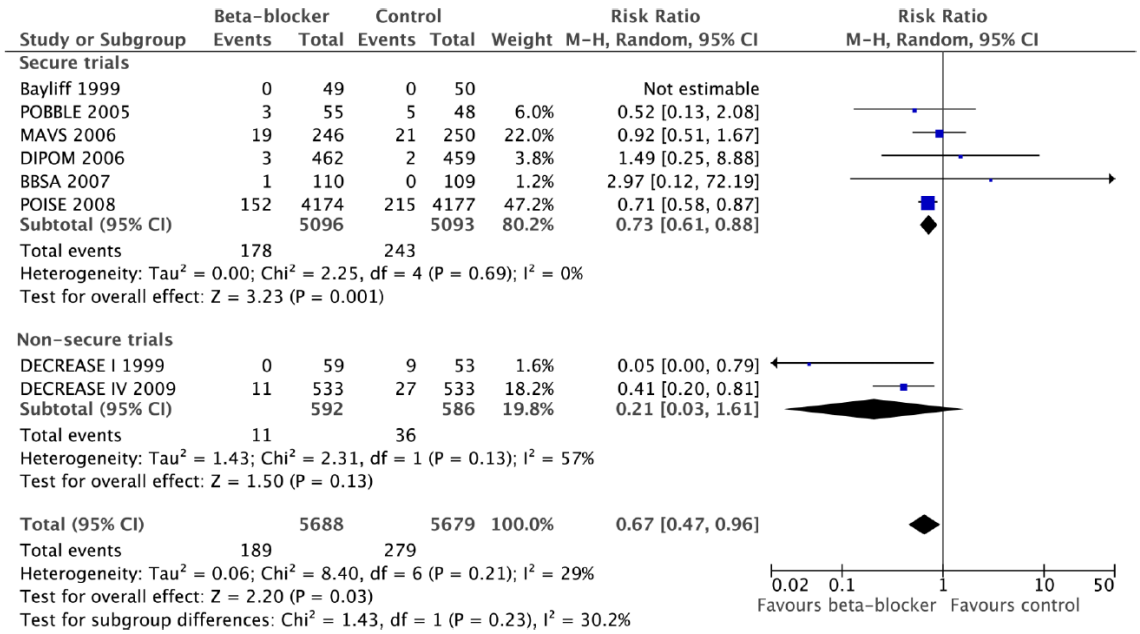


Figure 5 Comparison of effect of perioperative β-blockade on non-fatal myocardial infarction in secure and non-secure trials.

Critical appraisal of forest plots:

1. Is the data similar enough between studies to allow aggregation and meta-analysis? Visual inspection of the point estimates and the CI relative to the line of unity will inform this decision.
2. What is the magnitude of heterogeneity associated with the data, as reflected by the p-value and the I^2 statistic?
3. How precise are the results?

References

1. Peat J, Barton B. Medical statistics. A guide to data analysis and critical appraisal. Great Britain: Blackwell Publishing 2005.
2. Polgar S, Thomas SA. Introduction to research in the health sciences. Third ed. United Kingdom: Churchill Livingstone 1995.
3. Norman GR, Streiner DL. Biostatistics. The bare essentials. Third ed. Connecticut, USA: People's Medical Publishing House 2008.
4. Guyatt GH, Rennie D, Meade MO, et al. Users' guides to the medical literature: A manual for evidence-based clinical practice. Third ed. New York: McGraw-Hill Education 2015.
5. Older P, Smith R, Courtney P, et al. Preoperative evaluation of cardiac failure and ischemia in elderly patients by cardiopulmonary exercise testing. *Chest* 1993;104(3):701-4.
6. Cuthbertson BH, Card G, Croal BL, et al. The utility of B-type natriuretic peptide in predicting postoperative cardiac events and mortality in patients undergoing major emergency non-cardiac surgery. *Anaesthesia* 2007;62(9):875-81. doi: 10.1111/j.1365-2044.2007.05146.x
7. Le Manach Y, Hofer CK, Lehot JJ, et al. Can changes in arterial pressure be used to detect changes in cardiac output during volume expansion in the perioperative period? *Anesthesiology* 2012;117(6):1165-74. doi: 10.1097/ALN.0b013e318275561d [published Online First: 2012/11/09]
8. Suehiro K, Tanaka K, Yamada T, et al. The utility of intra-operative three-dimensional transoesophageal echocardiography for dynamic measurement of stroke volume. *Anaesthesia* 2015;70(2):150-9. doi: 10.1111/anae.12857 [published Online First: 2014/10/01]
9. Potgieter D, Simmers D, Ryan L, et al. N-terminal pro-B-type Natriuretic Peptides' Prognostic Utility Is Overestimated in Meta-analyses Using Study-specific Optimal Diagnostic Thresholds. *Anesthesiology* 2015;123(2):264-71. doi: 10.1097/ALN.0000000000000728
10. Choi JH, Cho DK, Song YB, et al. Preoperative NT-proBNP and CRP predict perioperative major cardiovascular events in non-cardiac surgery. *Heart* 2010;96(1):56-62. doi: 10.1136/hrt.2009.181388
11. Devereaux PJ, Yang H, Yusuf S, et al. Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): a randomised controlled trial. *Lancet* 2008;371(9627):1839-47. doi: 10.1016/S0140-6736(08)60601-7 S0140-6736(08)60601-7 [pii] [published Online First: 2008/05/16]
12. Bouri S, Shun-Shin MJ, Cole GD, et al. Meta-analysis of secure randomised controlled trials of beta-blockade to prevent perioperative death in non-cardiac surgery. *Heart* 2014;100(6):456-64. doi: 10.1136/heartjnl-2013-304262 [published Online First: 2013/08/02]

Pharmacokinetics of TIVA/TCI

Dr Graeme Wilson

Division of Paediatric Anaesthesia
UCT Department of Anaesthesia & Perioperative Medicine

Total intravenous anaesthesia (TIVA) can be defined as a technique of general anaesthesia using a combination of agents given solely by the intravenous route and in the absence of all inhalational agents including nitrous oxide. The availability of short-acting, potent hypnotic and analgesic drugs and modern target controlled infusion (TCI) systems make the process of administering TIVA now a practical and straightforward alternative to conventional inhalational anaesthesia. Rapid and precise titration of anaesthetic drugs to provide smooth onset and short, predictable drug offset is now feasible. This may be advantageous in the context of a cost-conscious health service moving towards ever shorter inpatient stay and more patients treated as day cases. Competency in TIVA is also vital for safe management of patients with contraindications for inhalational anaesthesia who require general anaesthesia.

The main objectives in the administration of TIVA are rapid loss of consciousness and lack of awareness during the operation; the level of anaesthesia and analgesia should closely follow the level of surgical stimulation to ensure haemodynamic stability; the drug effects should rapidly wear off at the end of the operation so that the patient has no residual sedation, no respiratory depression and no painful sensation from the surgical trauma. TIVA is particularly concerned with an understanding of the time course of drug effect, and subsequently the practitioner to have a sound understanding of the pharmacokinetics involved. Poor understanding of the pharmacokinetics underlying TIVA has caused accidental awareness as documented in the Fifth National Audit Project on accidental awareness during general anaesthesia (NAP5) report. In this report, there were 28 probable reports of awareness involving intravenous anaesthesia. In 21 of them total intravenous anaesthesia (TIVA) was used for induction and maintenance of anaesthesia. In these cases, the commonest cause of awareness was the administration of an inappropriately low dose infusion, usually as a fixed-rate infusion regime.

The pharmacokinetics of induction and maintenance of TIVA

The administration of a bolus dose of a drug results in a peak (central compartment) concentration, which then decreases rapidly with time as the drug is redistributed into the peripheral compartments. (Fig 1.)

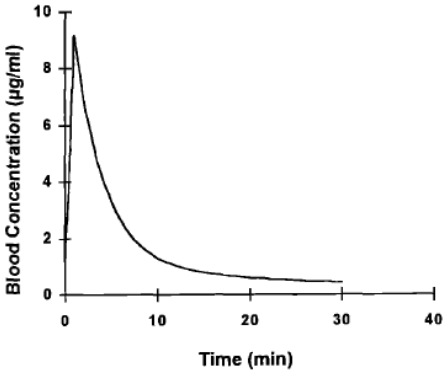


Fig 1 Plasma concentrations of propofol following a single bolus dose

Early drug distribution kinetics (front-end kinetics) determines the rate and extent of both drug distribution to the brain and its dilution by distribution to indifferent tissues. Following rapid injection there is a delay between the time drug is administered and the time drug appears at the sampling site. This is because drug is diluted by the venous return before entering the pulmonary circulation. The lungs delay the passage of drugs and even sequestrate significant amounts of drug. The systemic cardiac output then distributes drug to various organs and the drug becomes detectable in the arterial blood before being returned by the venous flow for recirculation. The exposure of the various organs to anaesthetic drug will be therefore determined by both the cardiac output as well as the distribution of the cardiac output. Both cardiac output and its distribution are important determinants of the early drug concentration versus time relationship of intravenously administered drugs and inter-individual variability in response to rapidly acting intravenous anaesthetics.

These plasma drug changes are best described by a three-compartment model (Fig 2.) The drug is administered to the central compartment 1 represented by the blood. It then redistributes into two peripheral compartments 2 and 3, the rate at which is determined by the rate constants K12 and K13 drug will also move in the opposite direction back into the central compartment as described by the rate constants K21 and K31. Clearance is described by the rate constant K10. The central nervous system, not the plasma is the effect site of hypnotic and analgesic drugs. There is a delay in their onset of action as the drugs diffuse into the CNS which can be described by the rate constant Ke0

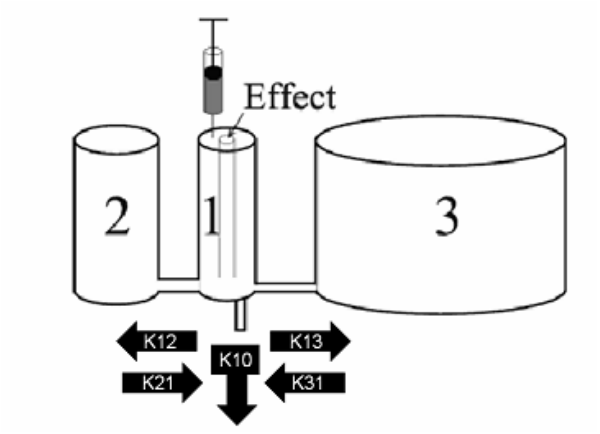


Fig2. Graphical representation of a three-compartment model

Although a rapid onset of action is desirable for the induction of anaesthesia with an intravenous agent, recovery of consciousness would be rapid without some form of maintenance. Repeated single bolus doses can be used to maintain a drug's effect, but it is easy to understand how such 'peaks and troughs' in drug concentration can result in both toxic and sub-therapeutic effects (Fig 3.) Another approach would be to simply run a constant infusion of drug. When drugs are given by infusion at a constant rate, a steady state drug concentration can be achieved, but to achieve such stability, a considerable period is required. Many of the drugs that we use have long elimination half-lives and steady state is approached only after the drug has been infused for approximately 5 elimination half-lives (Fig 4.) In terms of how long it will take to reach steady state, the answer is simple - infinity. The reason is that the steady state concentration is asymptotically approached, but never reached. For a propofol infusion this would mean that the plasma concentration would continue to slowly increase before approaching a steady state at somewhere around 24-36 hours.

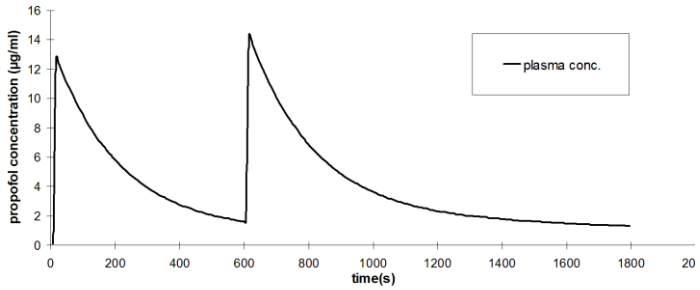
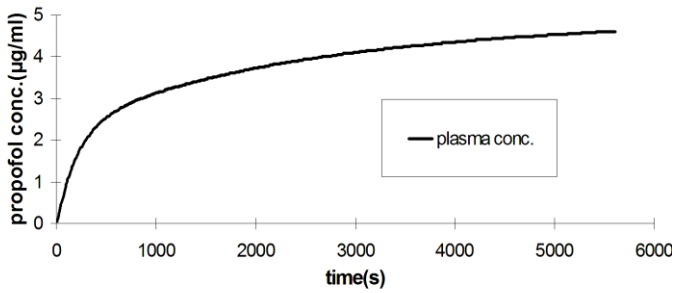


Fig 3. Plasma concentration of a drug following multiple boluses

Fig 4. Plasma concentration of a drug infused at a fixed rate.



Achieving a steady plasma concentration is the chief goal or aim when administering a TIVA. To achieve an almost constant plasma concentration in a drug whose pharmacokinetics are reliably represented by a three-compartment model, will require a loading dose and variable infusion rate.

In 1968 Kruger-Thiemer described a theoretical approach to achieving and maintaining a steady blood concentration of a drug. Briefly to achieve a steady state it is necessary to administer –

1) An initial bolus dose e to 'fill' the central compartment (and provide rapid induction of anaesthesia)

2) A constant final infusion rate maintaining central compartment concentration by matching the elimination of drug once redistribution is complete and drug concentrations in peripheral and central compartments have equilibrated

3) An interim infusion rate maintaining central compartment concentration by matching the rates of transfer of drug from central to peripheral compartments.

This became known as the BET scheme – Bolus Elimination Transfer scheme. This concept was instrumental in the development of the first TIVA recipes and TCI systems. In 1988, Roberts et al. proposed such a regimen for the delivery of propofol anaesthesia in adults. The manual infusion scheme for a target blood propofol concentration of 3 micrograms/ml, consisted of a loading dose of 1 mg/kg followed immediately by an infusion of 10 mg/kg/hour for 10 minutes, 8 mg/kg/hour for the next 10 minutes and 6 mg/kg/hour thereafter. An overall mean blood propofol concentration of 3.67 micrograms/ml was achieved within 2 minutes and maintained stable for the subsequent 80-90 minutes of surgery. Unfortunately, this regimen fails to provide adequate anaesthesia for all patients in all circumstances, risking excessive doses in some and proving inadequate in others. Subtle pharmacokinetic and pharmacodynamic variability between patients dictate the clinical effect of any given drug concentration and these differences, coupled with the need to vary depth of anaesthesia to combat changing levels of surgical stimulation, necessitate the ability to quickly and reliably titrate drug concentration to effect.

The pharmacokinetics of emergence from anaesthesia.

At the termination of a drug infusion plasma drug levels start to fall - this is due to clearance. Clearance is defined as the amount of plasma cleared of drug per unit time and consists of two main components namely redistribution and metabolic elimination of drug. As described above during an infusion drug is simultaneously metabolised and redistributed to peripheral compartments. Drug will continue to redistribute to peripheral compartments until a true steady state is achieved (after approximately 5 elimination half-lives) where clearance will be solely due to metabolism of drug. For most drugs, elimination occurs in an exponentially declining manner, the rate of elimination being proportional to the plasma concentration, as the downstream end of the gradient remains at zero. This system (i.e. the amount of drug being removed is a constant fraction in unit time rather than a constant amount) is known as first-order kinetics. As most of our anaesthetic agents usually follow first order kinetics clearance at true steady state will match the infusion rate.

Target controlled infusion (TCI)

TCI provides an accurate and user friendly method of delivering a TIVA. Alterations to the plasma or effect site concentration can be made rapidly and without the need for complex calculations and manual changes in the infusion rate. One feature of TCI pumps that contributes greatly to the anaesthetist's ability to accurately time the offset of drug effect is the ability of the pumps to calculate the "relevant effect-site decrement time". A typical TCI system consists of three main components: a user interface incorporating a display and method to input data, a microprocessor to run the pharmacokinetic modelling software and control the third component namely the infusion device. The infusion device needs to be able to deliver high infusion rates typically up to 1200ml/hour within a precision of 0.1ml/hour.

Achieving a steady plasma concentration with TCI

When a target concentration is selected the TCI-pump administers a bolus to rapidly fill the central compartment. The size of the bolus is calculated from the initial volume of distribution and if applicable the difference between the pre-existing compartment concentration and the target concentration.

When the pharmacokinetic model determines that the target concentration has been reached the infusion rate decreases. The new infusion rate will be determined by the calculated clearance of the drug and the redistribution of the drug to peripheral compartments. Eventually, when the peripheral compartments are saturated (i.e. at steady state) the infusion rate will match the clearance of the drug. If a target concentration less than the present blood concentration is selected, the TCI pump will stop the infusion until the target concentration is reached as estimated by the pharmacokinetic model used.

Pharmacokinetic models

There is a mathematical relationship between an administered dose of a drug and the resulting observed changes in plasma concentration. This relationship allows mathematical pharmacokinetic models to be constructed that may then be used to facilitate the calculation of dosing regimens and to guide pharmacotherapeutic management. A pharmacokinetic model is a mathematical model that predicts the plasma concentration of a drug after administration by infusion or bolus. The data for these models are derived by plasma concentration measurement in volunteers given a bolus or infusion of the drug being studied, and the model is derived using statistical techniques.

As the pharmacokinetics of different drugs varies, each drug will obviously need its own pharmacokinetic model. Similarly, when significant pharmacokinetic differences exist within the population, such as the difference between adults and children, different pharmacokinetic models will be needed.

Several pharmacokinetic models exist for adult and paediatric propofol TCI and several adult TCI models exist for the short acting opioids.

Choosing a TCI model

Most confusion arises when using propofol as the user is requested to choose between different pharmacokinetic models. The two main adult models in use are those described by Marsh and Schnider. The use of a weight-proportional pharmacokinetic model, (Marsh model), means that all the volumes and clearances of a multi-compartment model are proportional to only the weight of the patient. Interestingly, when using a weight-proportional pharmacokinetic model with a TCI pump, a doubling of the weight has the same effect on the infusion rates as a doubling of the targeted concentration. The Marsh model will adequately predict the plasma and effect site concentrations in most non-elderly adults who are of a normal body habitus. In the elderly, the Marsh model will typically display a wide variability. The Schnider model uses age and lean body mass as co-variates, and it may be a safer model to use when administering propofol to both elderly and overweight patients. The Ke0 values of the Marsh and Schnider result in differing times to peak effect (TTPE). The TTPE of Schnider's model is 1.6 min whereas the TTPE with the Marsh model is 4.57 min. When using Schnider's model for effect site targeting there is less plasma overshoot and better cardiovascular stability in compromised or elderly patients.

Two paediatric models are available for use at present in South Africa - one developed by Kataria et al and the second adapted from the preliminary models developed by Schüttler and incorporated in the Paedfusor software.

Kataria's model is based on a relatively small population (53 children) ranging from 3 to11 years. It uses both weight and age as covariates in determining infusion rates. There are no peer reviewed publications validating the accuracy of this model. The Paedfusor model is the newest paediatric model and has recently become commercially available in South Africa. The Paedfusor model is validated for children from 6 months (min 5 kg) to 16 years. A recent peer reviewed publication show that the Paedfusor model performed well in children undergoing cardiac surgery or cardiac catheterisation.

In practical terms the Kataria model has a smaller calculated volume of distribution and faster clearance and redistribution compared to the Paedfusor model. This equates to a smaller initial bolus and faster infusion rate when increasing the plasma target concentration.

At present the there is only one pharmacokinetic model available for remifentanil, sufentanil and alfentanil each on the commercially available TCI pumps. These are the Minto, Gepts and Maitre models respectively. None of these have been validated for use in children.

Effect site

So far we have focused on plasma drug concentration. This may be misleading, because the plasma is not the site of drug effect. For example, even though the plasma concentration following an intravenous bolus peaks nearly instantaneously after the bolus, no anaesthesiologist would induce a patient with an intravenous bolus of a hypnotic and immediately proceed with intubation. Figure 5 shows the time delay between plasma concentration and EEG effect of fentanyl and alfentanil, as reported by Scott and Stanski.

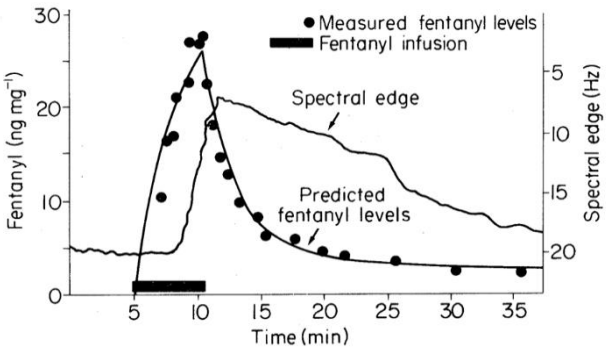


Fig. 5 the EEG effect is delayed nearly 3 minutes after the plasma concentrations of fentanyl rise. The EEG effect of alfentanil follows the rise in drug concentration much more closely, suggesting more rapid equilibration between the plasma and the site of drug effect.

This delay between peak plasma concentration and peak effect is called hysteresis. Hysteresis is the clinical manifestation of the fact that the plasma is not the site of drug action, only the mechanism of transport. Drugs exert their biological effect at the effect site, which is the immediate milieu where the drug acts upon the body, including membranes, receptors, and enzymes. The rate of equilibration between blood and effect site depends on several factors – cardiac output, cerebral blood flow, lipid solubility degree of ionisation etc.

The concentration of drug in the effect site cannot be measured. First, it is usually inaccessible, at least in human subjects. Second, even if we could take tissue samples, the drug concentration in the microscopic environment of the receptive molecules will not be the same as the concentration grossly measured in, say, ground brain or CSF. Although it is not possible to measure drug concentration in the effect site, using rapid measures of drug effect we can characterize the time course of drug effect. Knowing the time course of drug effect, we can characterize the rate of drug flow into and from the effect site. Knowing these rates, we can characterize the drug concentration in the effect site in terms of the steady state plasma concentration that would produce the same effect. Starting with the 3-compartment model we can incorporate the effect site as an additional compartment.

The effect site is the hypothetical compartment that relates the time course of plasma drug concentration to the time course of drug effect. ke0 is the rate constant of drug elimination from the effect site. The effect compartment receives such small amounts of drug from the central compartment that it has no influence on the plasma pharmacokinetics.

Effect site vs plasma targeting

The problem with targeting the plasma concentration is that when the target concentration is changed there is a temporal delay before the concentration at the effect site changes. When targeting the plasma concentration, the user determines the plasma concentration and the effect site concentration follows passively with a time delay determined by the Ke0 It is possible to target the effect site concentration using the Ke0 in combination with the other pharmacokinetic parameters. The TCI system then manipulates the blood concentration to change the effect site concentration as rapidly as possible. This will necessitate an “overshoot” in the plasma concentration to achieve a plasma-effect site gradient to cause rapid effect site equilibration. As it is the plasma concentration that determines the cardiovascular side effect profile care should be taken when effect site targeting in elderly or compromised patients.

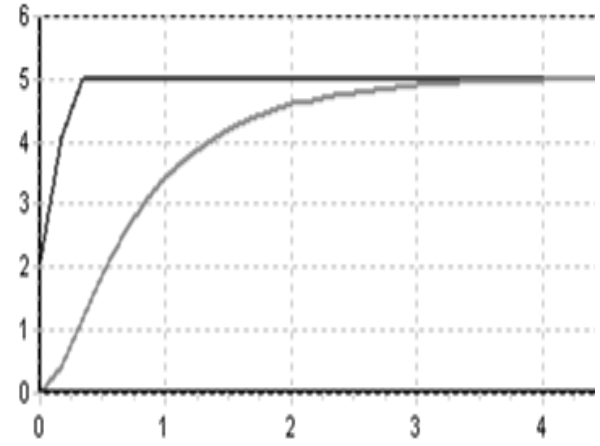


Fig 6. Plasma targeting. Targeting plasma concentration results in a slower induction

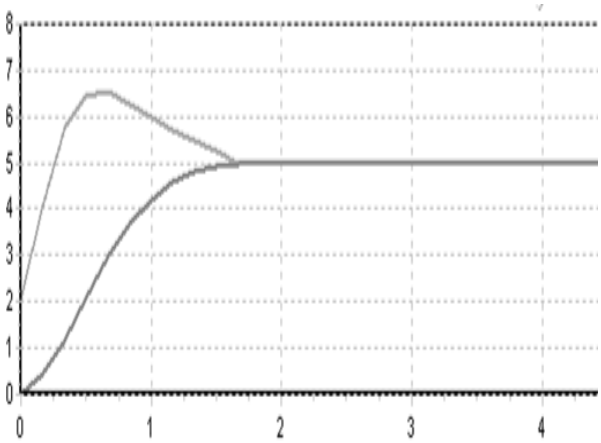


Fig 7. Effect site targeting. Targeting effect site results in overshoot of the plasma concentration

Conclusion

TCI and TIVA do not change the properties of drugs. Target concentrations are calculated, not measured. TCI pumps maintain three superimposed infusions, one at a constant rate to replace drug elimination and two exponentially decreasing infusions to match drug removed from central compartment to other peripheral compartments of distribution. Nowadays TCI technology is becoming a part of routine anaesthesia technique for the practitioner rather than a research tool for specialists and those who are enthusiasts of intravenous anaesthesia. Besides clinical application in anaesthesia, target controlled systems will play a significant role as research tools in the evaluation of drug interactions in anaesthesia and in the development of new control techniques for the administration of sedative and analgesic drugs in the peri-operative period.

FURTHER READING

Minto CF and Schnider TW. Contributions of PK/PD Modeling to Intravenous Anesthesia. Clinical Pharmacology & Therapeutics 2008 84, 27-38.
Absalom A, Struys MMRF. An overview of TCI & TIVA. 2nd edn. Gent: Academia Press, 2006.
Absalom AR, Mani V, De Smet T, Struys MM. Pharmacokinetic models for Propofol defining and illuminating the devil in the detail. Br J Anaesth 2009; 103: 26-37.
Struys MM, Sahinovic M, Lichtenbelt BJ, Vereecke HE, Absalom AR. Optimizing intravenous drug administration by applying pharmacokinetic/pharmacodynamic concepts. Br J Anaesth. 2011 Jul;107(1):38-47.
Vuyk J1, Mertens MJ, Olofsen E, Burm AG, Bovill JG. Propofol anesthesia and rational opioid selection: determination of optimal EC50-EC95 propofol-opioid concentrations that assure adequate anesthesia and a rapid return of consciousness. Anesthesiology. 1997 Dec;87(6):1549-62.
Roberts FL, Dixon J, Lewis GTR, Tackley RM, Prys-Roberts C. Induction and maintenance of propofol anaesthesia. Anaesthesia 1988; 43.
Marsh B, White M, Morton N, Kenny GN. Pharmacokinetic model driven infusion of propofol in children. Br J Anaesth 1991; 67.
Schnider TW, Minto CF, Gambus PL, et al. The influence of method of administration and covariates on the pharmacokinetics of propofol in adult volunteers. Anesthesiol 1998; 88.
Hughes, M.A., Glass, P.S. & Jacobs, J.R. Context-sensitive half-time in multicompartment pharmacokinetic models for intravenous anesthetic drugs. Anesthesiology 76, 334–341 (1992).
Roberts FL, Dixon J, Lewis GT, Tackley RM, Prys-Roberts C. Induction and maintenance of propofol anaesthesia. A manual infusion scheme. Anaesthesia 1988 Mar;43(Suppl.):14–7

Cytochrome P450

Dr Brigid Brennan

UCT Department of Anaesthesia & Perioperative Medicine

The cytochrome P450 (CYP450) enzymes are a major determinant of the pharmacokinetic behavior of numerous drugs.

CYP450 enzymes are so named because they are bound to membranes within the cell, specifically the endoplasmic reticulum (cyto) and contain a heme pigment (chrome and P) that absorbs light at a wavelength of 450nm when exposed to carbon monoxide. Each cytochrome P450 isoenzyme consists of a single protein chain and one haem group as the binding site for the drug.

The CYP450 enzymes are found predominantly in the liver but also exist in the small intestine (reducing drug bioavailability), brain, lung, adrenal gland, kidney, bone marrow, skin, ovary, testes, and placenta.

Classification

There are about 50 different CYP's found in humans.
There are many different isoforms of cytochrome P450. An isoform is a CYP enzyme variant that derives from one particular gene. These isoenzymes are classified according to similarities of their amino acid sequencing into families (number), subfamilies (letter) and individual genes/specific enzymes (number).

Families: Members of a family must have at least 40% sequence homology. Families are numbered e.g. CYP 1, CYP 2. There are at least 74 CYP families but only about 17 have been described in man.

Subfamilies: Members of a subfamily must have at least 55% sequence homology. About 30 subfamilies are well described in humans. Subfamilies are identified by a letter e.g. CYP2D

Individual genes: There are about 50 important genes in man. Individual genes are identified by a number e.g. CYP2D6

CYP450 and metabolism

Metabolism involves enzymatic conversion of one chemical entity to another within the body. Biotransformation is the metabolism of substances foreign to the body.

Most drugs are highly lipid soluble and nonpolar and are not easily eliminated by the kidney so most lipophilic substances are converted to more polar/hydrophilic products which can then be excreted in the urine. Biotransformation makes drugs water soluble.

Most drugs are metabolized by two groups of enzymes; CYP450 and non-microsomal enzymes (e.g. monoamine oxidase, alcohol dehydrogenase). The cytochrome P450 enzyme system catalyses the metabolism of endogenous and exogenous compounds. Here we are predominantly interested in the catabolism of drugs by the cytochrome P450 system.

Drug metabolism occurs predominantly (but not exclusively) in the liver. Other sites include kidney, lungs, blood, GIT and placenta. Drug metabolism involves two kinds of reactions; Phase 1 and Phase 2. These usually, although not always, occur sequentially.

Phase 1 reaction: Metabolic transformation.
Catabolic reaction that occurs by oxidation, reduction or hydrolysis. The result of this transformation can be:

Active drug converted to inactive metabolite (most common).
Inactive drug converted to active metabolite (prodrug e.g. tramadol).
Active drug converted to active metabolite (prolonging the effect of the drug e.g. morphine).

Phase 1 reactions often introduce a reactive group, such as hydroxyl, into the molecule. This then serves as the point of attack for the conjugating system to attach a substituent e.g. glucuronide. CYP450 is responsible for 70-80% of phase 1 metabolism. Oxidation is the most common phase 1 reaction. Phase 1 oxidative reactions catalyzed by CYP450 include epoxidation, N-dealkylation, O-dealkylation, S-oxidation and hydroxylation. Oxidation can result in both activation and inactivation of a compound.

Drug oxidation by the P450 system requires drug/substrate, P450 enzyme, molecular oxygen, NADPH and a flavoprotein (NADPH P450 reductase). Basically one atom of oxygen is added to the drug to create a hydroxyl group, the other atom of oxygen is converted to water. A typical cytochrome P450 oxidation reaction is as follows, where R is the substrate/drug:

$$\mathrm{NADPH + H^+ + O_2 + RH \Rightarrow NADP^+ + H_2O + R\text{-}OH}$$

Reductive reactions are much less common than oxidation reactions (e.g. warfarin inactivated by conversion of a ketone to a hydroxyl group by CYP 2A6)

Phase 2 reaction: Conjugation reaction
Phase 2 reactions are synthetic/anabolic. Involves conjugation which usually results in an inactive product. A covalent link is formed between a functional group on the substrate and acetate, amino acids, glucoronic acid or glutathione. These polar inactive products can then be excreted in the urine/faeces. Phase 2 metabolism does not require cytochrome P450 enzymes.

Function of CYP450 enzymes

CYP is integral to the metabolism of both endogenous and exogenous compounds.

Physiological function:

- Biosynthesis and degradation of endogenous compounds e.g. steroid hormones, cholesterol and fatty acids.
- Physiological role in the brain where it is involved in the release of peptide hormones from the hypothalamus and pituitary.
- Regulation of vascular tone in the brain by arachidonic acid metabolites.
- CYP2D6 may regulate metabolism and processing of neurotransmitters such as dopamine and serotonin and have a role in determining mental state/personality.
- Vascular autoregulation particularly in the brain.
- Adrenal and gonadal steroidogenesis are influenced by the CYP450 enzymes.
- CYP450 has a role in the control of body fluid volume and composition and hence blood pressure, by its action on arachidonic acid, the metabolites of which have both vasoconstrictive and vasodilatory activity.

Pharmacological function:

- CYP450 in the liver plays a major role in drug metabolism by converting drugs from a hydrophobic state to a more readily excretable hydrophilic form, mainly by an oxidation reaction. Metabolism of exogenous substances (drugs, environmental pollutants, chemicals, anaesthetic agents) may produce metabolites that are toxic or carcinogenic.
- CYP450 enzymes are essential for the production of cholesterol, steroids, prostacyclins and thromboxane A2.
- Detoxification of foreign chemicals/toxins.

CYP450 enzymes

These enzymes differ from one another in amino acid sequence, in sensitivity to inhibitors and inducers, and in the specificity of the reactions that they catalyse. Different members of the family have distinct but often overlapping substrate specificities with some enzymes acting on the same substrates but at different rates. One drug can be metabolized by several different isoenzymes.

There are 74 CYP gene families, but CYP families 1-3 are responsible for the metabolism of drugs in the human liver.

CYP450's in the CYP2C, 2D and 3A subfamilies are most active in metabolizing known clinically used drugs. CYP3A4 and CYP2D6 metabolise most anaesthetic drugs.

The CYP's most important in oxidative metabolism are:

- CYP3A4
- CYP2D6*
- CYP2C9*
- CYP2C19*
- CYP2A6*
- CYP2E1
- CYP1A2

* polymorphic

Enzyme induction can occur with all of these, except CYP2D6.
The following are polymorphic: CYP2D6, CYP2C9, CYP2C19, and CYP2A6.
More than one CYP450 can be involved in the metabolism of a particular drug.

Specific enzymes

CYP3A4: These enzymes account for 30% of liver P450 and 70% of gut P450. CYP3A4 is the most common cytochrome subfamily in the liver. Significant first pass metabolism of midazolam is brought about by mucosal CYP3A4 of the small intestine. Rifampicin is the most potent inducer of CYP3A4. The CYP3A4 enzyme is responsible for the metabolism of several drugs used in anaesthesia: Opioids (fentanyl, sufentanyl, alfentanil), benzodiazepines (midazolam, diazepam) and local anaesthetics (lignocaine). One study showed that the elimination half-life of midazolam was prolonged by 50% and its clearance reduced by 30% by the co-administration of fentanyl at induction, probably as a result of competitive inhibition of CYP3A4 activity.

CYP2D6: This enzyme has been extensively studied. Genetic polymorphisms of this enzyme exist. The enzyme is NOT inducible by pharmacological agents. Drugs that inhibit the enzyme essentially turn the patients taking them into poor metabolisers. This enzyme is responsible for the biotransformation of codeine to morphine. Substrates: Codeine, beta blockers, antidepressants, antipsychotics, neuroleptics, antiarrhythmic. Anaesthetic drugs: Codeine, tramadol, ondansetron, granisetron.

CYP2C9: Substrates include NSAIDS, warfarin.

CYP2C19: Substrates include TCA's, diazepam, clopidogrel.

CYP2E1: Responsible for metabolism of volatiles. CYP2E1 is a major catalyst for the formation of trifluoroacetylated proteins from halothane which have been implicated as target antigens responsible for halothane hepatitis.

CYP1A2 metabolises ropivacaine

Polymorphisms, enzyme induction and enzyme inhibition

Within the human population, there are major sources of interindividual variation in P450 enzymes. Drug-drug interactions can occur when one drug alters the metabolism of another drug by inhibition or induction of the CYP450 enzymes. Metabolic enzymes are under genetic control and so may vary between different populations or individuals.

Changes in activity of CYP450 can therefore result from:

- Genetic polymorphism
- Enzyme inhibition
- Enzyme induction
- Physiological/environmental factors (vitamin deficiencies, pregnancy, fasting)

These factors are of major importance in therapeutics.

Genetic polymorphism

In different people and different populations, **activity** of CYP450 oxidases differs. Because of this different people may respond differently to the same drug. Genetic variation in a population is termed polymorphism when both gene variants exist with a frequency of at least 1%.

Polymorphisms are relevant if:

1. The metabolising enzyme is responsible for 50% or more of the clearance of the drug.
2. The drug has a steep dose-response curve and a narrow therapeutic window.
3. Drug activity depends on an active metabolite formed by a polymorphic enzyme.

Individuals can be classed as having poor, intermediate, extensive (normal) or ultra-rapid CYP450 activity.

A specific gene encodes each CYP450 enzyme. Every person inherits one genetic allele from each parent. Alleles are 'wild type' or 'variant'. 'Wild type' occurs most commonly in the general population. Variant alleles encode an enzyme with reduced or no activity. The frequency of variant alleles varies significantly and depends on race and ethnic background.

Poor metaboliser: 2 copies of variant alleles (lack both copies of the functional allele).

Intermediate metaboliser: Heterozygous for one functional (wild type) and one deficient (variant) allele or two partially defective alleles that cause reduced enzyme activity.

Extensive (normal) metaboliser: 2 copies of wild type alleles (functional allele). Majority of population.

Ultra-rapid metaboliser: Multiple copies of wild type alleles (3 or more functional alleles) resulting in excess enzyme activity.

Genetic variations in CYP450 metabolism should be considered when patients exhibit unusual sensitivity or resistance to drug effects at normal doses.

Testing is available to categorise CYP2D6 metabolism as poor, intermediate, extensive and ultra-rapid. This can be helpful in evaluating the efficacy and dosing of many medications used in breast cancer treatment and psychiatry. Although testing may be beneficial in anaesthesia, associated adjustments in dosing and clear recommendations have not been fully developed and have not become standard practice.

Examples of polymorphisms

CYP2D6 is polymorphic. Codeine's analgesic activity is as a result of its conversion to morphine by CYP2D6. In CYP2D6 poor metabolisers codeine will be ineffective as an analgesic. Ultra-rapid metabolisers will produce more morphine from a standard codeine dose and put the patient at risk of an opioid overdose. This has been implicated in infant toxicity from breastfeeding in mothers taking codeine analgesia. Other drugs metabolized by CYP2D6 e.g. antidepressants and neuroleptics, in poor metabolisers, will predispose the patient to drug toxicity. B-blocker removal can be impaired in 2D6 poor metabolisers.

Another polymorphism with significant consequences is deficient activity of CYP2C9. These patients will be ineffective in clearing (S)-warfarin…so could be fully anticoagulated on just 0.5mg warfarin a day.

Clopidogrel is a prodrug that requires biotransformation to an active metabolite by cytochrome p450 enzymes for its antiplatelet effect. Patients with reduced CYP2C19 activity have significantly lower levels of the active metabolite of clopidogrel, diminished platelet inhibition and therefore a higher rate of adverse cardiovascular events including stent thrombosis. CYP2C19 function is absent in 30% of Chinese people and 15% Caucasian. Genotyping is recommended by the AHA and ACC for those at high cardiovascular risk being treated with clopidogrel.

One beneficial CYP450 phenotype can be seen in patients deficient in CYP2C19…they have higher cure rates for peptic ulcer disease treated with omeprazole due to sustained high plasma levels achieved.

Does polymorphism affect drug design? Designing and making drugs costs a fortune. The interaction between CYP450 and newly designed drugs is so important to pharmaceutical companies that predominant degradation of a drug by one of the polymorphic CYP's is often enough to stop any further research/development on that drug!

CYP polymorphisms may be implicated in an individuals' susceptibility to disease e.g. lung cancer and Parkinson's disease.

Enzyme induction and enzyme inhibition

Knowledge of the CYP450 system and its substrates is a key factor in the prevention of important drug-drug interactions, either due to enzyme induction or inhibition. Drugs interact with the CYP450 system in several ways. A drug may be metabolized by only one enzyme or by multiple enzymes (warfarin CYP1A2, 2D6, 3A4). Drugs that cause metabolic drug interactions are referred to as either inducers or inhibitors.

Enzyme induction

An important cause of drug interactions and adverse drug reactions.

Certain drugs induce the enzymes responsible for their metabolism (rifampicin, phenobarbitone). These enzymes also metabolise other drugs. The clearance of the 2nd drug will therefore be increased by an enzyme inducer. This increased clearance means that a greater dose of the 2nd drug will be required to have the same therapeutic effect as that seen without enzyme induction.

Inducers increase CYP450 activity by increasing enzyme synthesis (DNA transcription) so there is usually a delay before enzyme activity increases. Most of the CYP450's can be induced except 2D6! Over 200 drugs can cause enzyme induction.

Result of enzyme induction:

Decreases concentration of the substrate drug. Increased dose required for same therapeutic effect.

Prodrug: Increased production of active drug metabolite. Patient at risk of overdose.

A drug can be metabolized by the same enzyme that it induces. In other words, the inducing agent is normally itself a substrate for the induced enzyme so this can result in slowly developing tolerance. Pharmacokinetic tolerance is less marked than pharmacodynamic tolerance (opioids) but is clinically important when starting treatment with carbamazepine. It is an enzyme inducer but is started at a low dose to avoid toxicity as enzymes are not yet induced, and then gradually increased over a few weeks during which time it induces its own metabolism and its half-life gradually decreases over time.

Rifampicin is the most notable enzyme inducer. Three days of rifampicin treatment will decrease the effectiveness of warfarin.

CYP3A4 metabolises many substrates and is induced by rifampicin, carbamazepine, phenytoin and dexamethasone. This will increase the metabolism of opioids, benzodiazepines and local anaesthetics.

If the active metabolite of a drug is toxic e.g. NAPQI metabolite of paracetamol, then enzyme induction will increase the risk of toxic side effects. The risk of serious hepatic injury following paracetamol overdose is increased in patients whose CYP450 system has been induced, e.g. by the

chronic use of alcohol.

Enzyme inhibition
Inhibitors block the metabolic activity of CYP450 enzymes.

The extent to which an inhibitor affects metabolism of a drug depends on factors like the dose of the inhibitor and the ability of the inhibitor to bind to the enzyme. A drug can be metabolised by and inhibit the same enzyme e.g. erythromycin, or it can be metabolised by one enzyme and inhibit another. Inhibitory effects usually occur immediately.

Inhibition take place by a variety of mechanisms:

- Drug may compete for active site but is itself not a substrate.
- Non-competitive inhibitors bind to the enzyme.
- Mechanism based inhibitors require oxidation by a P450 enzyme, oxidation products bind to haem.
- Direct irreversible inactivation

Result of inhibition of CYP450:
Increase the concentration of the substrate drug.

Decreased activity of prodrug metabolized by enzyme e.g. Clopidogrel is a prodrug, requiring metabolism by CYP450 into its active form to prevent arterial thrombosis. Omeprazole (PPI) inhibits the metabolism of clopidogrel and so reduces the efficacy of the drug in preventing arterial thrombosis. Other PPI's e.g. lansoprazole, have a much lower affinity for the CYP enzyme and so do not affect the efficacy of clopidogrel.
Propofol interferes with the metabolism of alfentanil and sufentanil by inhibiting CYP2B1 and CYP1A1 and so may potentially alter the metabolism of co-administered alfentanil/sufentanil.

Severe toxicity can result if enzyme inhibiting drugs are added to the following medications: Atypical antipsychotics, benzodiazepines, statins, warfarin.

Common inhibitors: Cimetidine, azole antifungals (ketoconazole), HIV protease inhibitors, calcium channel blockers, erythromycin, SSRI antidepressants. Cimetidine, an H2 receptor antagonist, inhibits many different CYP's and is now rarely used because of its inhibitory effect on CYP450.

Drugs can be intentionally combined to take advantage of CYP450 inhibition: Ritonavir, a protease inhibitor and potent CYP3A4 inhibitor, is added to lopinavir to boost serum levels of the later in HIV.

Detailed information of metabolizing enzymes and of drug interactions for common drugs is increasingly available in texts or electronically.

Environmental / Physiological
Dietary chemicals can affect the concentration of CYP enzymes by a variety of mechanisms such as changes in the rate of gene transcription or degradation of mRNA. Cigarette smoke is an enzyme inducer. Fasting has been shown to induce CYP2E1 enzyme. In general vitamin deficiencies lower CYP450 activity.

Table 1. Significant Cytochrome P450 Enzymes and Their Inhibitors, Inducers, and Substrates

Enzyme	*Potent inhibitors**	*Potent inducers†*	*Substrates*
CYP1A2	Amiodarone (Cordarone), cimetidine (Tagamet), ciprofloxacin (Cipro), fluvoxamine (Luvox‡)	Carbamazepine (Tegretol), phenobarbital, rifampin (Rifadin), tobacco	Caffeine, clozapine (Clozaril), theophylline
CYP2C9	Amiodarone, fluconazole (Diflucan), fluoxetine (Prozac), metronidazole (Flagyl), ritonavir (Norvir), trimethoprim/sulfamethoxazole (Bactrim, Septra)	Carbamazepine, phenobarbital, phenytoin (Dilantin), rifampin	Carvedilol (Coreg), celecoxib (Celebrex), glipizide (Glucotrol), ibuprofen (Motrin), irbesartan (Avapro), losartan (Cozaar)
CYP2C19	Fluvoxamine, isoniazid (INH), ritonavir	Carbamazepine, phenytoin, rifampin	Omeprazole (Prilosec), phenobarbital, phenytoin
CYP2D6	Amiodarone, cimetidine, diphenhydramine (Benadryl), fluoxetine, paroxetine (Paxil), quinidine, ritonavir, terbinafine (Lamisil)	No significant inducers	Amitriptyline, carvedilol, codeine, donepezil (Aricept), haloperidol (Haldol), metoprolol (Lopressor), paroxetine, risperidone (Risperdal), tramadol (Ultram)
CYP3A4 and CYP3A5	Clarithromycin (Biaxin), diltiazem (Cardizem), erythromycin, grapefruit juice, itraconazole (Sporanox), ketoconazole (Nizoral), nefazodone (Serzone‡), ritonavir, telithromycin (Ketek), verapamil (Calan)	Carbamazepine, *Hypericum perforatum* (St. John's wort), phenobarbital, phenytoin, rifampin	Alprazolam (Xanax), amlodipine (Norvasc), atorvastatin (Lipitor), cyclosporine (Sandimmune), diazepam (Valium), estradiol (Estrace), simvastatin (Zocor), sildenafil (Viagra), verapamil, zolpidem (Ambien)

CYP=cytochrome P.

**—These will slow down substrate drug metabolism and increase drug effect.*
†—These will speed up substrate drug metabolism and decrease drug effect.
‡—Brand not available in the United States.

Am Fam Physician 2007

Cancer and CYP450

There is some speculation about the role of CYP proteins and polymorphisms as causes of cancer, some may activate pro-carcinogens to carcinogens and many are involved in the removal of carcinogens from the body. For cancers that are hormone sensitive, CYPs involved in steroid metabolism may play a role in suppression/promotion of malignancies through such metabolism. CYP2D6 has been implicated in mediating carcinogenesis by activating procarcinogens in tobacco smoke leading to lung cancer. Genetic polymorphisms of CYP2E1 may play a role in the development of hepatic cancer.

CYP and age

There is variable expression of CYP enzymes in different people, different race groups and at different ages. CYP 1A2 is not expressed in neonates making them particularly susceptible to toxicity from drugs such as caffeine. Fetuses and neonates have low activity of CYP2D6 and can functionally be considered 2D6 poor metabolisers. CYP3A4 activity is also low in term and preterm neonates, increases in early childhood and then gradually decreases to adult levels around puberty. In the elderly, an age-related 20% reduction in metabolism of CYP2D6 substrates has been observed but the same has not been seen for the CYP3A family.

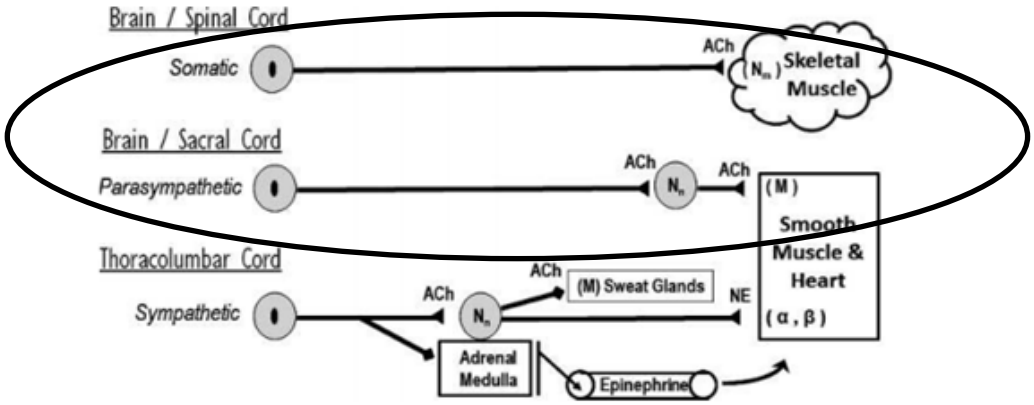
References

1. Milner A, Welch E. Applied pharmacology in Anaesthesiology and Critical Care. Chapters 1, 3, 4.
2. Vaja R, Barker R. Drugs and the liver. Anaesthesia and Intensive Care Medicine 2012; 13:2. 71-74.
3. Dale O, Olkkola K.T. Cytochrome P450, molecular biology and anaesthesia. Editorial. Acta Anaesthesiol Scand 1998; 42: 1025-1027.
4. Seefelder C, Leeder J.S. Cytochrome P450 pharmacogenetics and anaesthesia. Correspondence. Paediatric Anaesthesia 2002; 12: 809-812.
5. Chang GWM, Kam PCA. The physiological and pharmacological roles of cytochrome P450 isoenzymes. Anaesthesia 1999; 54: 42-50.
6. Charlton M, Thompsons J. Adverse drug reactions. Anaesthesia and Intensive Care Medicine 2017; 18:4. 205-209.
7. Mega JL et al. Cytochrome P450 Polymorphisms and response to Clopidogrel. The New England Journal of Medicine 2009; 360; 4: 354 – 361.
8. Lynch T, Price A. The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects. Am Fam Physician 2007; 76(3): 391-396.

The Parasympathetic Nervous System and Associated Pharmacology

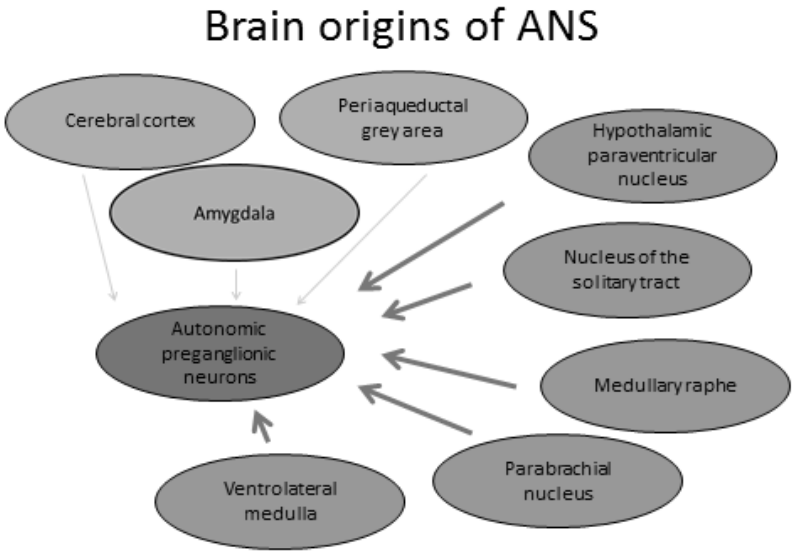
Dr Janieke van Nugteren

UCT Dept of Anaesthesia & Perioperative Medicine
Pain Management Unit



Becker DE. Basic and Clinical Pharmacology of Autonomic Drugs. *Anesthesia Progress*. 2012;59(4):159-169.

Pathways from the brain that control autonomic responses.

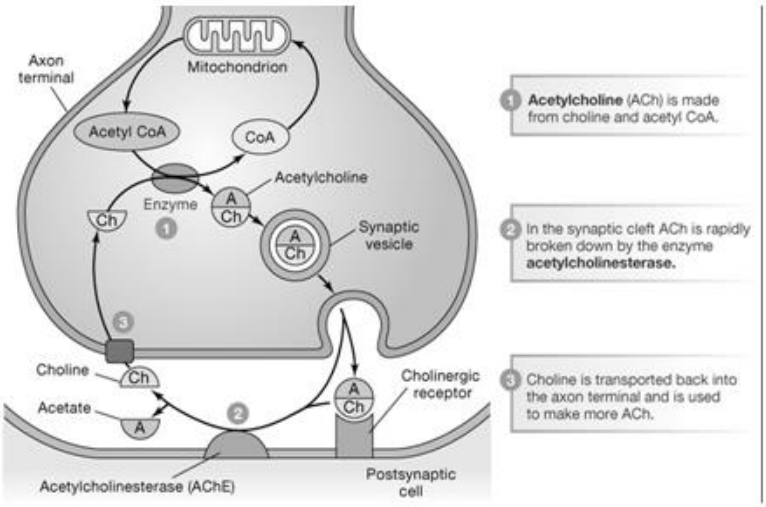


Indirect projections (thin arrows) Direct projections (thick arrows)

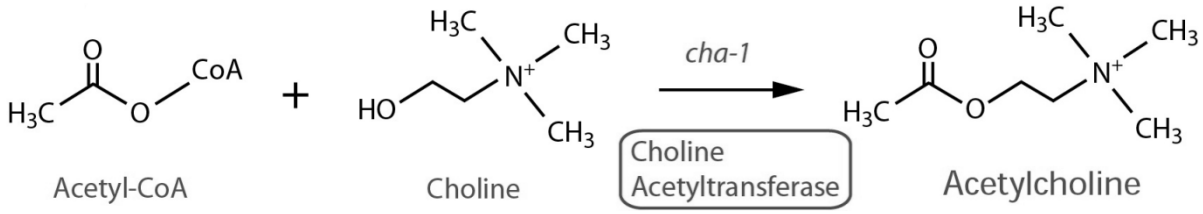
The parasympathetic nervous division of the autonomic nervous system (ANS) is sometimes called the craniosacral division of the ANS because of the location of the preganglionic neurons. The parasympathetic nerves supply the visceral structures in the head via the oculomotor, facial and glossopharyngeal nerves – and those in the thorax and upper abdomen via the vagus.
The sacral outflow supplies the pelvic viscera via branches of the 2^{nd}-4^{th} sacral spinal nerves.
Parasympathetic preganglionic fibres synapse on ganglia clustered within wall of the visceral organs. This means that the parasympathetic postganglionic fibres are very short.

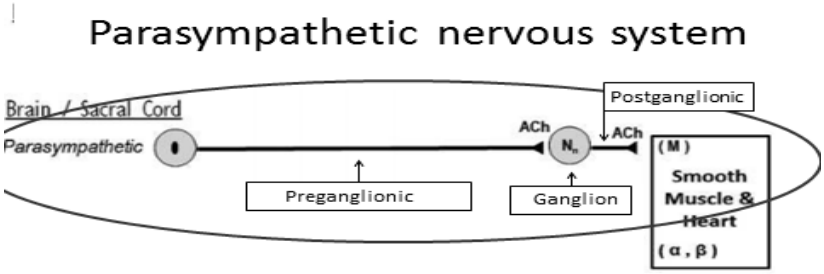
Neurotransmitters:

Acetylcholine is released at all preganglionic neurons, postganglionic parasympathetic neurons and a few post-ganglionic sympathetic neurons (sympathetic vasodilatory fibres and sweat glands).
The remaining sympathetic postganglionic neurons release **norepinephrine.**
Acetylcholine is formed within cholinergic neurons and broken down within the synaptic cleft:



Cholinergic neuron

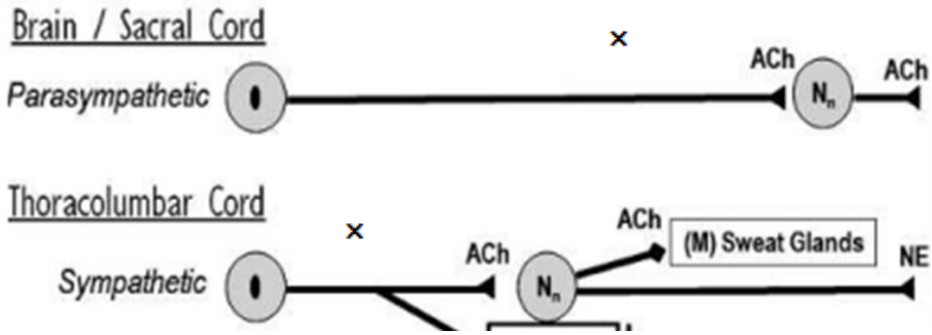




Ganglion agonists		Ganglion antagonists	
Nicotine	Stopping smoking	Hexamethonium	(First anti-hypertensive)
Lobeline		Trimetaphan	BP lowering
Epibatidine			

Ganglionic transmission is mediated via **nicotinic agonists** and blocked via **nicotinic antagonists.**
None of these drugs other than nicotine (which is used in smoking cessation products) are used clinically anymore.
Ganglion antagonists block both SNS and PNS ganglia and so have the following clinical effects:

- Hypotension / ↓ HR
- Loss of cardiac reflexes
- Inhibition of secretions
- GIT paralysis
- Impaired micturition



Muscarinic agonists (Direct)

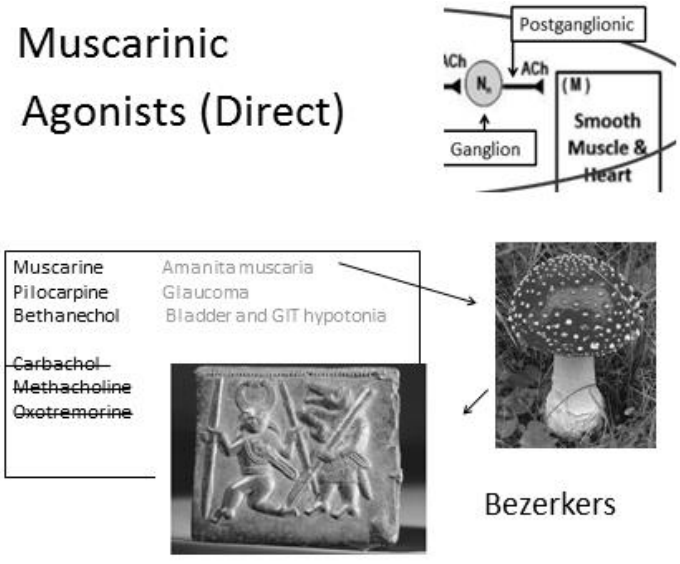
Muscarine was first isolated from the mushroom Amanita muscaria.

Anaesthetists do not use drugs in this class.

The majority of drugs which are used in this class are those that are used to treat glaucoma. The parasympathetic nerves to the eye supply the constrictor pupillae muscle which runs circumferentially in the eye. The ciliary muscle adjusts the curvature of the lens. In glaucoma drainage of aqueous humour becomes impeded with a dilated pupil. Activation of the constrictor pupillae muscles by muscarinic agonists lowers IOP.

Pilocarpine is a tertiary amine and crosses the conjunctival membrane and is used to treat glaucoma.

Bethanechol stimulates smooth muscle of the GIT and facilitates bladder evacuation in the absence of organic obstruction.



Muscarinic agonists (Indirect)

Anticholinesterases inhibit acetylcholinesterase (AChE) -the enzyme that causes hydrolysis of Ach to choline and acetic acid. Anticholinesterases inhibit hydrolysis by binding to the AChE and forming relatively stable complexes to prevent Ach reaching the catalytic site of the enzyme.
They can be divided into:

- Reversible anticholinesterases
- Irreversible anticholinesterases – Organophosphates

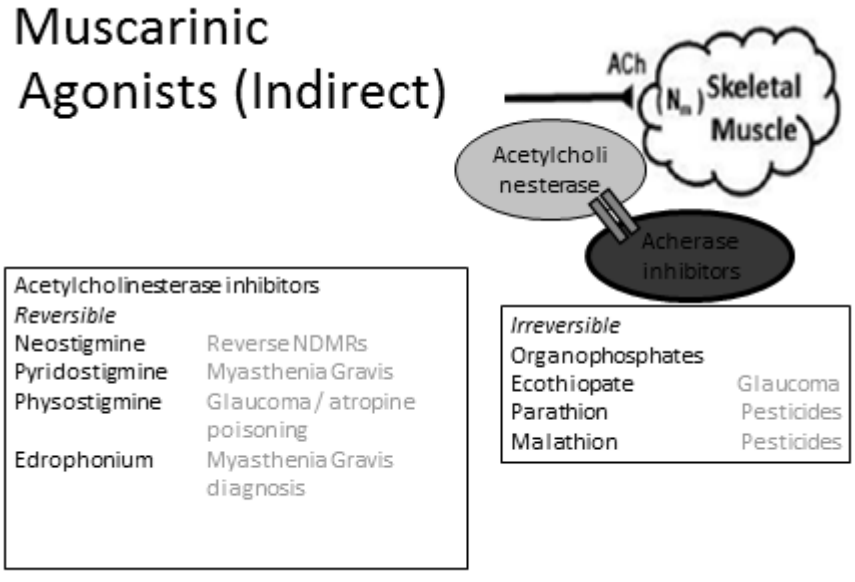
Reversible anticholinesterases

	Edrophonium	Physostigmine	Neostigmine	Pyridostigmine
Indications	Tensilon test: In MG helps to distinguish cholinergic crisis (too much Ach) and too little (Myasthenic crisis)	-Glaucoma -Postop shivering	-Reversal of NDMRs -Myasthenia Gravis	Myasthenia Gravis (drug of choice due to long DOA/ less muscarinic side effects)
Structure	Quaternary ammonium	Tertiary amine (crosses the BBB)	Quaternary ammonium	Quaternary ammonium
Predominant site of action	Presynaptic	Postsynaptic	Postsynaptic	Postsynaptic
Onset of action	80 seconds	Not known	120 seconds works at same time as glycopyrrolate	3-5 minutes(very slow)
Duration	60-120 min after IV	20-30 min	40-60 min IV 2-4 hours after oral	112 minutes

Adapted from Applied Pharmacology in Anaesthesiology and Critical Care: Multi author, Milner A, Welch EH 1st edition 2012

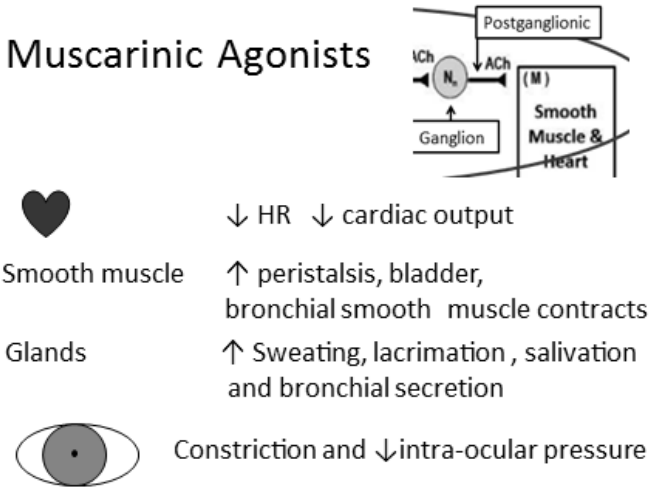
Irreversible anticholinesterases

Organophosphate poisoning is relevant to anaesthetists since these patients present to ICU.
Examples of these pesticides are Parathion and Malathion.



They combine with AChE at the esteratic site to form a stable phosphorylated inactive complex which cannot be hydrolysed.
Treatment in ICU involves atropine infusion which acts as a muscarinic antagonist.

Echothiopate is used by ophthalmologists to cause ciliary muscle contraction for treatment of glaucoma.
It may enhance neostigmine's effect and should not be used prior to GA with neuromuscular blocking agents

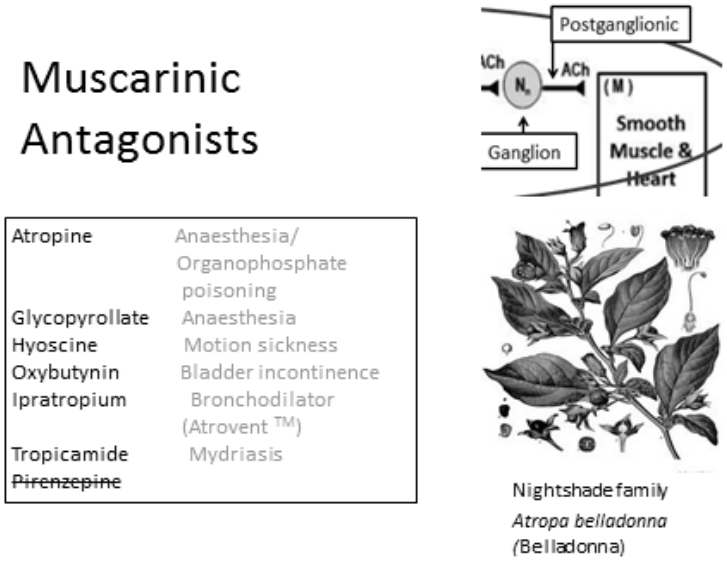


Muscarinic antagonists

Atropine and hyoscine are alkaloids from the belladonna plants (Nightshade family– which includes tomatoes and aubergine). They are muscarinic antagonists – resulting in competitive antagonism of muscarinic receptors- thus blocking the effect of Ach. In general they don't cause significant effects at the nicotinic receptors.

Atropine and hyoscine are tertiary amines which cross the blood brain barrier (BBB).

Ipratropium (Atrovent ™) used to treat bronchospasm and glycopyrrolate are quaternary amines and do not cross the BBB.



Red as a beet (Hyperthermia)

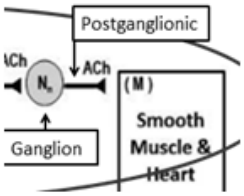
Dry as a bone (No sweating)

Mad as a hatter (crosses BBB)

Blind as a stone (mydriasis and poor near-sightedness).

Ways to remember the effects of atropine

Muscarinic Antagonists



↑HR ↑cardiac output

Smooth muscle — Decreased peristalsis, bladder, bronchial smooth muscle dilation

Glands — ↓ Sweating, lacrimation , salivation and bronchial secretion

Dilation and ↑ intra-ocular pressure

	Atropine	Glycopyrrolate
Onset	50-80 secs	90-120 secs
Duration	2-4 hours	Vagal block 2-3 hours Antisialogogue 7 hours (IM)
Dose	10-20 mcg/kg IV	4-6 mcg/kg
Characteristics	Tertiary amine- crosses BBB	Quaternary amine- does not cross BBB More powerful antisialogogue
Side effects	Dry mouth Relaxes LOS Relaxes bronchi (Inspissation of secretions) Confusion (elderly) Decreased sweating/ heatstroke	Dry mouth Relaxes LOS
Caution	IHD / High temp/ pyrexia/ obstructive uropathy/ glaucoma	Glaucoma/ CVS disease/ High temp/ pyrexia

The Oxygen-Haemoglobin Dissociation Curve

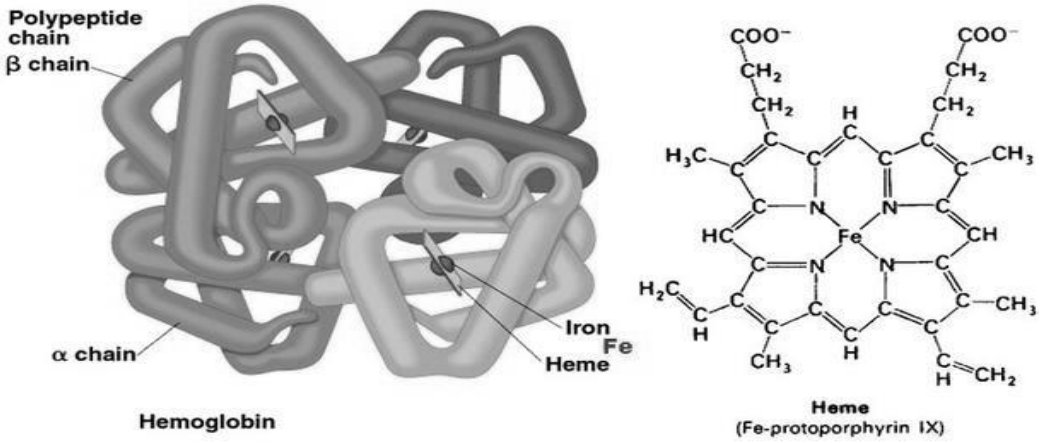
Dr Alma de Vaal

UCT Dept of Anaesthesia & Perioperative Medicine

Haemoglobin and the transport of O_2

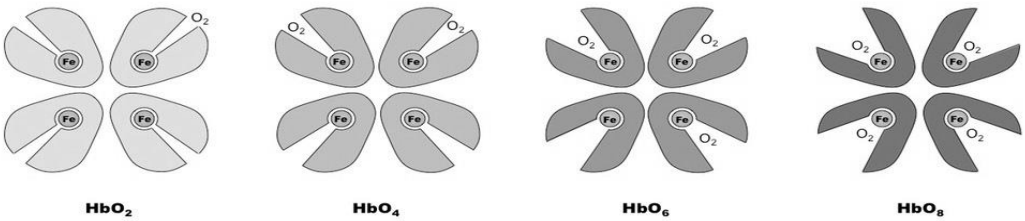
Oxygen is transported across the alveolar-capillary membrane into the bloodstream. Most (97%) of the O2 binds to Haemoglobin (Hb) in the red blood cells, however a small amount (3%) dissolves in the plasma. Oxygen is then transported to the peripheral tissues where it moves out of the bloodstream and gets utilized as fuel for oxidative cellular metabolism.

Haemoglobin forms part of the red blood cell and consists of two parts, haem and globin. Succinyl-CoA, formed in the Krebs cycle, binds with glycine to form a pyrrole molecule. Four (4) pyrroles combine to form a porphyrin ring that combines with iron in the ferrous state (Fe^{2+}) to form haem. Haem is conjugated with globin, a polypeptide that consists of four subunits. The most common form of haemoglobin is HbA that comprises 2-alpha and 2- beta polypeptides.



O_2 binds reversibly to haem, each haemoglobin molecule can bind up to four O_2 molecules. Hb is an allosteric protein; the binding of oxygen to one haem group increases the affinity within the remaining haem groups. The tetrameric globular structure of hemoglobin is adapted to our physiological needs to regulate oxygen delivery far better than primitive globins such as muscle myoglobin and cytoglobin.

When O_2 binds to haemoglobin the iron remains in the ferrous state and the reaction is called oxygenation. The haem-haem interaction is a function of the tetrameric nature of haemoglobin implies that 4 haem groups do not undergo simultaneous oxygenation or deoxygenation.



This is called positive 'cooperativity' – each oxygen molecule binds more easily than the previous one. Oxy-haemoglobin has a substantially different quaternary structure than deoxyhaemoglobin.
When O_2 binds to haem, it pulls the haem molecule closer towards the plane of the proto-porphyrin ring, flattening and changing its shape. Ionic interactions holding the 4 globin chains together are distracted and as they reposition and the quaternary structure is altered, which increases the binding affinity of the other globin chains. When fully oxygenated with 4 O_2 molecules the haemoglobin is in a 'R' or a relaxed state whereas in fully deoxygenated haemoglobin, the quaternary structure, is in a 'T' or a tense state making it difficult for oxygen to gain access to haem.

A few important concepts

Oxygen content: Arterial oxygen content (CaO_2) is the amount of oxygen bound to haemoglobin plus the amount of oxygen dissolved in arterial blood:

CaO_2 (ml O_2/dl) = (1.39 x haemoglobin concentration x SaO_2) + (0.0031 x PaO_2)

SaO_2 is the arterial oxy-haemoglobin saturation, **PaO_2** is the arterial oxygen tension and 1.39 is the **Hüfner** constant – explained below.

Normal CaO_2 is approximately 20 mL O_2/dL.

The Hüfner Constant

Haemoglobin oxygen capacity is expressed by the maximum volume of oxygen that combines with 1g of haemoglobin - known as the Hüfner constant.
A theoretical value based on the molecular weight of haemoglobin which is 64458g and has been calculated as 1.39ml.g^{-1}. *In vivo* experiments, however, produce values slightly lower based on the presence of small amounts of other forms of haemoglobin which are relatively poor carriers of oxygen, such as methaemoglobin, carboxyhaemoglobin and HbA_2.

(22414ml x 4)/64458g = 1.39ml

The value of 22414ml is the volume of 1 mole of ideal gas (Avogadro's law) at STP (standard temperature and pressure) 0˚ Celsius, 760mmHg. The volume of 1 mole of gas at body temperature (37 ˚) will be larger but haemoglobin's oxygen capacity is less than the calculated value and the clinical value generally accepted varies from 1,31 – 1,39 millilitres.

When haemoglobin is 100% saturated, each gram of normal Hb contains 1.39ml of O_2. When the Hb in normal blood is around 15g/dL then 100ml of blood contains 20.1ml (1,39 ml x 15) of O_2 bound to Hb when Hb is 100% saturated. The amount of dissolved O_2 is a linear function of the PO_2.
In vivo the blood at the ends of the pulmonary capillaries is only 97% saturated with a PO_2 of 97mmHg because of the slight admixture of venous blood that bypasses pulmonary capillaries (physiological shunt). Systemic arterial blood is then only 97% saturated and therefore only contains 19.8ml of O_2 per 100ml blood. (0.29ml in solution and 19.5ml bound to Hb).
When blood passes through tissue capillaries this amount reduces to 14.4ml of O_2 per 100ml blood (PO2 of 40mmHg and 75% saturated haemoglobin). Therefore, under normal conditions, 5ml of oxygen is transported to the tissues by each 100ml of blood.

During exercise the transport of oxygen changes. The muscle cells use more O_2 and in extreme cases O_2 may fall to levels as low as 15mmHg with only 4.8ml of oxygen remaining. Then 19.8 - 4.8 = 15mmHg which is the quantity of oxygen transported by each 100ml of blood, there are three times more O_2 transported in each volume that passes through normally. Cardiac output also increases 6-7 times which gives a 20-fold increase in O_2 delivery to the tissues. The percentage of blood that gives up O_2 as it passes through the tissue capillaries is called the utilization coefficient and normally about 25%. With strenuous exercise this value can increase to 75-85%.

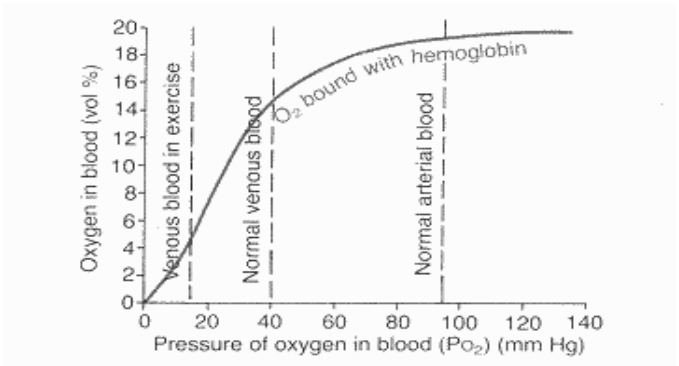
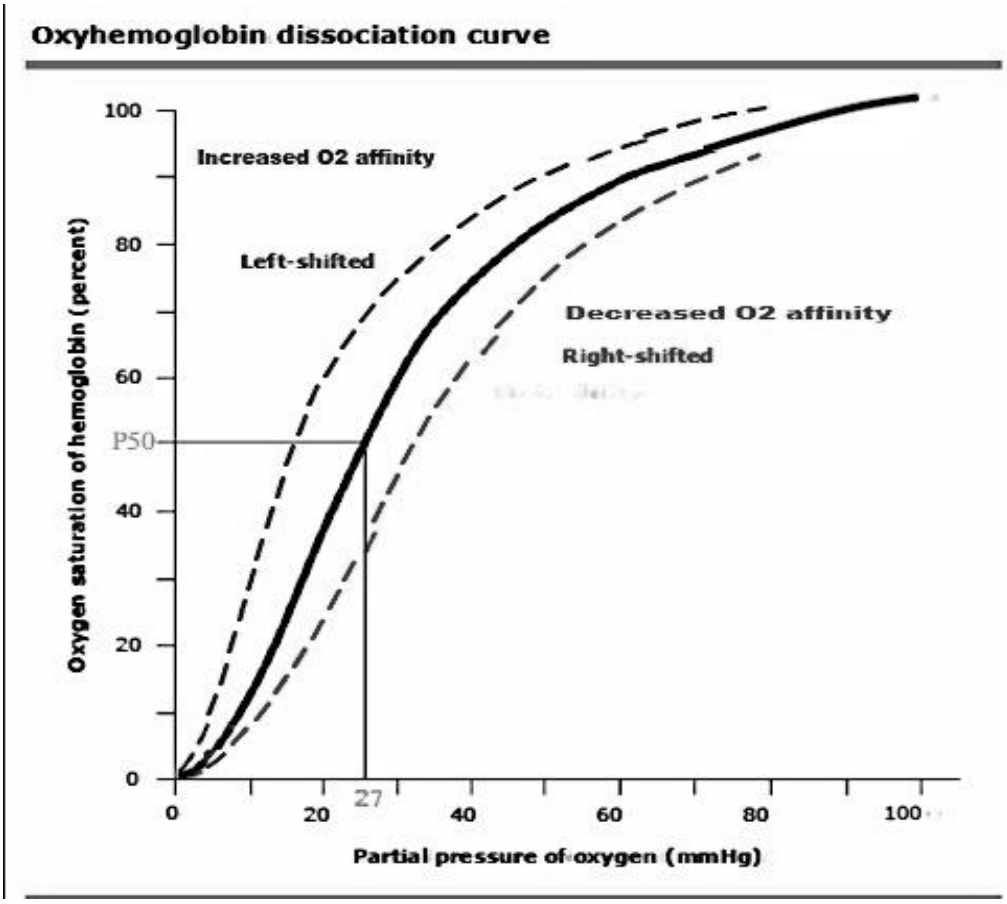


Figure 40–9. Effect of blood P_{O_2} on the quantity of oxygen bound with hemoglobin in each 100 milliliters of blood.

The relationship between the saturation of haemoglobin and the partial pressure of oxygen in the blood



This is the curve that relates the partial pressure of oxygen PaO_2 on the X-axis to the saturation of oxygen (percentage of haemoglobin saturated with oxygen) on the Y-axis.

The Sigmoid shape of the curve is because haemoglobin's affinity for oxygen changes as each sequential molecule of oxygen binds, in other words 'co-operativity'. The affinity of haemoglobin for oxygen is the lowest when the first molecule of O_2 binds to the tense deoxyhaemoglobin molecule. At a very low partial pressure of oxygen (PO_2), the gradient of the curve is almost flat. As the partial pressure of oxygen increases, more oxygen will bind and the curve gradient will increase until the maximum that can bind is reached. Very little additional binding occurs after that and the curve levels out as Hb is saturated with O_2.

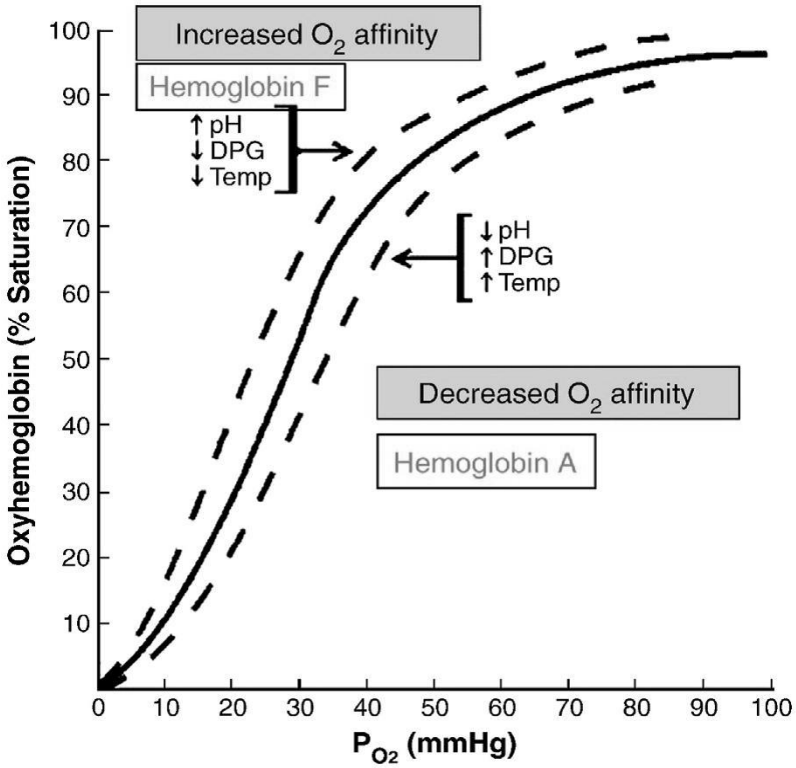
The amount of oxygen bound to haemoglobin at any time is related to the partial pressure of oxygen to which the haemoglobin is exposed. In the lungs at the alveolar-capillary interface, where the partial pressure of O_2 is high, the forward reaction is favoured. As blood circulates to peripheral tissues where the partial pressure of oxygen is less, the reverse reaction reducing haemoglobin, is favoured and haemoglobin will release oxygen into the tissue. Both reactions are extremely rapid taking less than 0,01seconds.

As blood in the arteries has a PO_2 of about 95mmHg the usual oxygen saturation of arterial blood is about 97% and normal venous blood returning from the tissues has a PO_2 of about 40mmHg and the saturation of haemoglobin is about 75%.

The P_{5O} is the partial pressure of oxygen in the blood in which haemoglobin is 50% saturated. This is an important concept and used as a measure of oxygen affinity which is used to compare the changes of the position of the curve. P_{50} may vary according to physical and chemical factors which may shift the curve to the left or right. Different forms of haemoglobin also have different P_{50} values. The higher the P_{50}, the lower the affinity of haemoglobin for O_2. Normal P50 is 27mmHg at 37˚C and a pH of 7,4.

Mixed venous O2 tension is approximately 40mmHg and the oxygen saturation of haemoglobin is 75%.

Factors that will affect the oxygen affinity of haemoglobin



A. Temperature
B. Hydrogen ion concentration (pH)
C. Carbon dioxide (pCO2)
D. 2,3 Diphosphoglycerate (2,3 DPG)

Temperature
A rise in temperature will shift the oxy-haemoglobin curve to the right. This means a higher PO_2 is required for haemoglobin to bind to a given amount of O_2.
A fall in temperature will shift the curve to the left and a lower PO_2 is required to bind a given amount of O_2.

pH (Hydrogen ion concentration)
When the pH is decreased, there will be a reduction in the oxygen's affinity for haemoglobin and the curve will move to the right. In an acidotic state, with tissue hypoxia, oxygen release will be enhanced to facilitate metabolically challenged areas. If the pH increases the alkalosis will move the curve to the left and there will be an increase in haemoglobin's affinity for oxygen. The Bohr effect will be covered in another lecture.

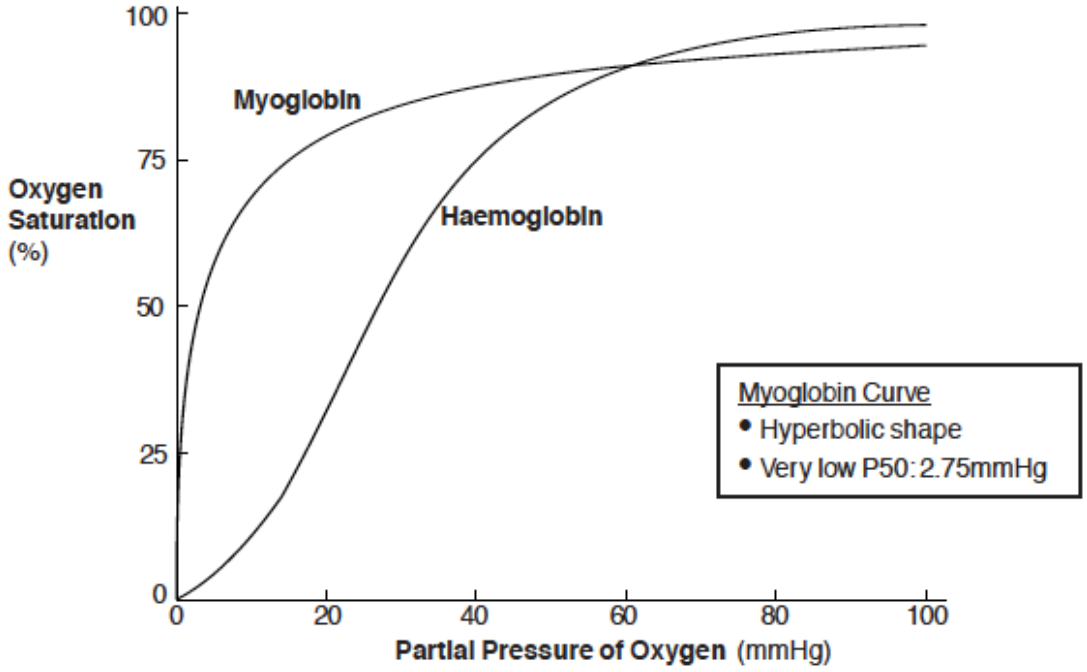
CO_2
Carbon dioxide is transported in the blood in 3 forms, most of which is in the form of bicarbonate ions (89%), secondly as carbamino compounds in reaction with haemoglobin (6%), and the rest is dissolved to form a solution. An increase in PCO_2 (and H+ ion concentration) will reduce the molecule's affinity for oxygen, thereby enhancing the release of oxygen from haemoblobin. This will also move the curve to the right.
The Haldane effect will be covered in another lecture.

2,3 DPG – 2,3 Diphosphoglycerate is plentiful in red blood cells and is formed as a product of glycolysis. It is a highly anionic organic phosphate and it promotes the release of oxygen from haemoglobin. One molecule of 2,3-DPG binds between the beta-globin chains of deoxy-haemoglobin thereby altering the protein structure and reducing oxygen affinity. It moves the curve to the right thus encouraging the unloading of oxygen at relatively low partial pressures of oxygen. When hypoxic conditions last longer than a couple of hours the quantity of 2,3-DPG in the blood will increase. The oxy-haemoglobin dissociation curve will shift even further to the right. It remains an important mechanism of adaptation during hypoxia.
The production of 2,3-DPG is increased in anaemia and with hypobaric hypoxia occurring at altitude. This results in more oxygen release to the tissues and moves the curve even further to the right. At

higher altitude, this beneficial effect will be opposed by a respiratory alkalosis secondary to hyperventilation that will move the curve to the left.
Interestingly stored blood contains 2,3 DPG that is metabolized whereby the ability of blood to deliver oxygen is reduced. The oxy-haemoglobin dissociation curve is moved to the far left. Blood still remains a better oxygen carrier than no blood at all but transfused red cells need more than 24 hours in the recipient before normal 2,3-DPG levels are restored.

Fetal haemoglobin (HbF) has a higher affinity for oxygen and therefore shifts the oxy-haemoglobin dissociation curve to the left. It has a reduced affinity for 2,3-DPG and the P50 is lower than normal adult haemoglobin. This is a type of Hb present in the fetus before birth that differs from normal adult haemoglobin. HbF increases oxygen release to fetal tissues under hypoxic conditions in which the fetus exists.

Myoglobin – is a haem containing oxygen binding protein present in the skeletal muscle and exhibits no co-operativity because it only contains a single globin chain, therefore the myoglobin dissociation curve lies to the far left and has the shape of a rectangular hyperbole. Myoglobin has a very low P_{50} (2,75mmHg) and the reason for this is because it needs to load and unload oxygen with PO_2 values that are in keeping with intracellular PO_2 levels.



Conclusion

Haemoglobin is exquisitely suited to carry oxygen optimally and maximize oxygen delivery to the tissues. The conformational changes that haemoglobin undergoes are perfectly reflected in the shape of the oxygen-haemoglobin dissociation curve which makes it a more effective carrier of oxygen when compared to other haemoglobin molecules or even oxygen dissolved in blood. Factors affecting the shape and position of the curve are important in terms of the affinity of oxygen to haemoglobin and are likely adaptive mechanisms to ensure maximal delivery of oxygen to the tissues in times of increased need.

References

1. Caroline Thomas, Andrew B Lumb. Physiology of haemoglobin, Continuous Education in Anaesthesia and Critical Care Pain, (2012) 12 (5): 251-256
2. William F Ganong, Review of Medical Physiology 18th edition
3. Guyton & Hall, Textbook of Medical Physiology 9thedition
4. UpToDate
5. Goralov V, Theoretiical value of Huffner's constant, Anaesthesia, (2004)59: 88-99

Metabolic Adaptation to Overnight Starvation

Dr Guy Picken

UCT Dept of Anaesthesia & Perioperative Medicine

Background

The first anaesthetic death occurred in 1848, two years after the first general anaesthetic was performed. Fifteen-year-old Hannah Greener aspirated and died while under chloroform anaesthesia for a toenail removal. Following this tragedy, it was proposed that a prolonged period of fasting should precede anaesthesia to allow the stomach to be empty of its contents. This tradition of overnight fasting became one of the most well-known routines in medicine. It was only challenged in the late 1980s and was found to be based on very little scientific evidence, and proved to be no safer than allowing patients to drink clear fluids up until 2 hours preoperatively. When the anaesthesia guidelines changed, follow-up investigations were performed and found no increase in complications (aspiration) and improved patient wellbeing (discomfort of thirst). More recently, Enhanced Recovery After Surgery (ERAS) takes the preoperative fasting guidelines one step further by advocating for a carbohydrate-rich drink 2 hours before induction of anaesthesia with the aim of reducing postoperative insulin resistance and protein catabolism.

Metabolic Effects of Overnight Fasting

The human organism requires a constant supply of fuel to provide the energy to survive. This energy intake is supplied intermittently as meals over the course of the day, while during the night we naturally fast.
The body contains these fuel stores:

- Carbohydrate (400g) – glycogen in liver & muscle – lasts 12-24 hours
- Protein (10-12kgs, only 20-30% is available as fuel) – lasts +- 21 days
- Fat (10-15kgs) – lasts +- 55 days

The post absorptive state (after the last glucose from the previous meal has been completely absorbed) occurs during the overnight fast when the effect of insulin fades away and glucose begins to be released from the liver (and lesser extent kidney) by glycogenolysis (glycogen breakdown) and gluconeogenesis (glucose synthesis from non-carbohydrate sources – lactate, alanine [muscle] and glycerol [fat]). The stimulus for this glucose production is a lack of substrates, resulting in the glucose level lowering and a resultant decreased insulin/glucagon ratio as the insulin level drops.

Should food be withheld for a longer period, there is transition from the ‘fasted state’ to starvation metabolism. This process (discussed in more detail later) is supported by several other hormones, in addition to low basal levels of insulin, like glucagon and cortisol, but also through complex interactions with growth hormone and the insulin-like-growth factor 1 system (IGF-1). IGF-1 is an anabolic hormone with effects on glucose and protein metabolism like those of insulin. Even brief fasting or hypocaloric nutrition for a few days (a regimen often used after abdominal surgery) results in marked reduction in insulin sensitivity (insulin resistance). The reduction in insulin levels and the low insulin/glucagon ratio is followed by reduced IGF-1 activity by an increase in IGF binding protein-1 (IGFBP-1).

Should a meal (or carbohydrate load) then be absorbed, there is a rapid switch from glycogenolysis to glycogenesis (glycogen synthesis) and increased peripheral glucose uptake (mainly in muscle) due to increased insulin levels and insulin/glucagon ratio. The delivery of lactate from anaerobic glucose metabolism maintains gluconeogenesis, although glucose-6-phosphate during these conditions is directed into liver glycogen (the indirect pathway). This results in almost completely inhibited glucose release from the liver and a rapid switch from net catabolism to net anabolism. In addition, levels of IGFBP-1 decrease, and IGF-1 activity increases.

Starvation Metabolism

Absolute starvation ultimately leads to death. Metabolic adaption to starvation is the body’s way of delaying this endpoint for as long as possible and allows us to survive starvation for up to 2 months. Since liver glycogen is depleted within 24 hours, gluconeogenesis (at the expenses of muscle protein)

becomes the sole provider of glucose for tissues that depend on glucose for survival – i.e. the brain, renal medulla and erythrocytes. During prolonged starvation, metabolism adapts in order preserve the protein stores from rapid depletion. This process is characterized by a reduction in T3 levels which leads to a decrease in the metabolic rate and muscle proteolysis. Reduced insulin levels along with activation of the adrenergic system result in an increased lipolysis in adipose tissue so that fats become the principle energy source. Glycerol from triglycerides (TGs) enters the glycolytic pathway and free fatty acids (FFAs) are broken down to acetyl-coenzyme A (acetyl-CoA). Excess acetyl-CoA results in formation of ketone bodies (ketosis). Some FFAs can contribute to gluconeogenesis. During prolonged starvation, the brain, kidneys and muscle begin to utilize ketone bodies efficiently as a source of fuel. The liver, through gluconeogenesis, can effectively recycle lactate (from glycolysis in glucose-dependent tissues) which is driven by energy derived from FFA oxidation and minimizes protein oxidation.

When FFAs and ketones are available as fuel, important mechanisms for protein-sparing are initiated. This occurs through the inhibition of key enzymes by the high acetyl-CoA/CoA ratio - 1) reduced glycolysis by pyruvate dehydrogenase inhibition & 2) reduced oxidation of amino acids by inhibition of branch chain 2-oxo-acid dehydrogenase. In starvation, the kidney becomes a gluconeogenesis tissue by converting the excess hydrogen ions that the kidney excretes as ammonia (from glutamine and glutamate) into glucose.

In summary, about 10-15% of the body weight can be lost without severe functional derangement, while weight loss of 35-40% becomes life-threatening.

Metabolic Effects from Surgery

Humans respond to surgery (and trauma) with multiple neuroendocrine changes leading to catabolism of stored body fuels and retention of salt and water. This '*surgical stress response'* was first described in the late 1920s, among patients admitted to hospital with long bone fractures. They found dramatic increases in nitrogen, potassium, phosphate, sulphur and creatine urinary losses and concluded that these represented a systemic breakdown in skeletal muscle. Later experimental studies showed increased levels of adrenal cortical hormones in response to injury. Furthermore, severing afferent nerve pathways from the site of injury diminished this response.

Modern understanding of the surgical stress response is that it involves activation of the sympathetic nervous system, secretion of catabolic hormones (cortisol, glucagon, growth hormone and catecholamines) and local cytokine responses to tissue injury. This response is usually proportional to the degree of surgical trauma or injury incurred. The endocrine component includes activation of the hypothalamic-pituitary-adrenal axis with increased cortisol secretion, increased secretion of vasopressin and increased pancreatic secretion of glucagon; this response leads to a net increase in peripheral insulin resistance and catabolism of skeletal muscle. The degree of peripheral insulin resistance has been linked to the magnitude of the catabolic response. In contrast to the situation during fasting, the metabolic adaption to minimize loss of body mass (particularly muscle mass) does not occur. Furthermore, injured tissues and tissues where synthesis of acute-phase proteins occurs still rely on glucose as a substrate. In addition, increasing body metabolism results in more severe net loss of fat and protein stores, and these responses are not to any great extent influenced by the availability of exogenous substrates.

The stress response to surgery has likely developed as an evolutionary response, allowing injured animals to survive without food and with healing of their wounds. However, in the current highly controlled surgical environment, this response is associated with several deleterious effects, including organ dysfunction, hypercoagulation, immunosuppression, catabolism and impaired wound healing. Peripheral insulin resistance is associated with hyperglycaemia - a possible cause of postoperative complications and an independent predictor of length of hospital stay.

Metabolic Effects of Preoperative Carbohydrate Drink

Only recently has it been shown that intake of clear fluids up to 2 hours before surgery may be permitted without an increase in gastric residual volumes or risk of aspiration of gastric contents. This prompted anaesthetic guidelines to allow free intake of clear fluids up to 2 hours prior to induction of anaesthesia. However, free fluids do not provide what is required to change the fasted state and reduce postop insulin resistance. Preoperative carbohydrate treatment aims to replicate normal

metabolic responses to eating breakfast. This treatment stimulates an endogenous insulin release, which switches off the overnight fasting metabolic state and is given to decrease the extent of peripheral insulin resistance while ameliorating the surgical stress response.

To test the hypothesis, a series of randomized studies were performed (Figure 1) where glucose was administered either as a preoperative infusion or as a carbohydrate-rich drink (400 mL, 50 g glucose) taken 2-3 hours before surgery. Scintigraphic studies showed that gastric emptying was complete within 2 hours after intake of this beverage. The amount of energy was enough to increase insulin to levels to that seen after a mixed meal, and insulin action enhanced by about 50% was shown 2-3 hours after intake. Randomized studies involving either preoperative glucose infusion or the carbohydrate-rich drinks all show that postoperative insulin resistance may be reduced by about 50% when preoperative fasting is avoided.

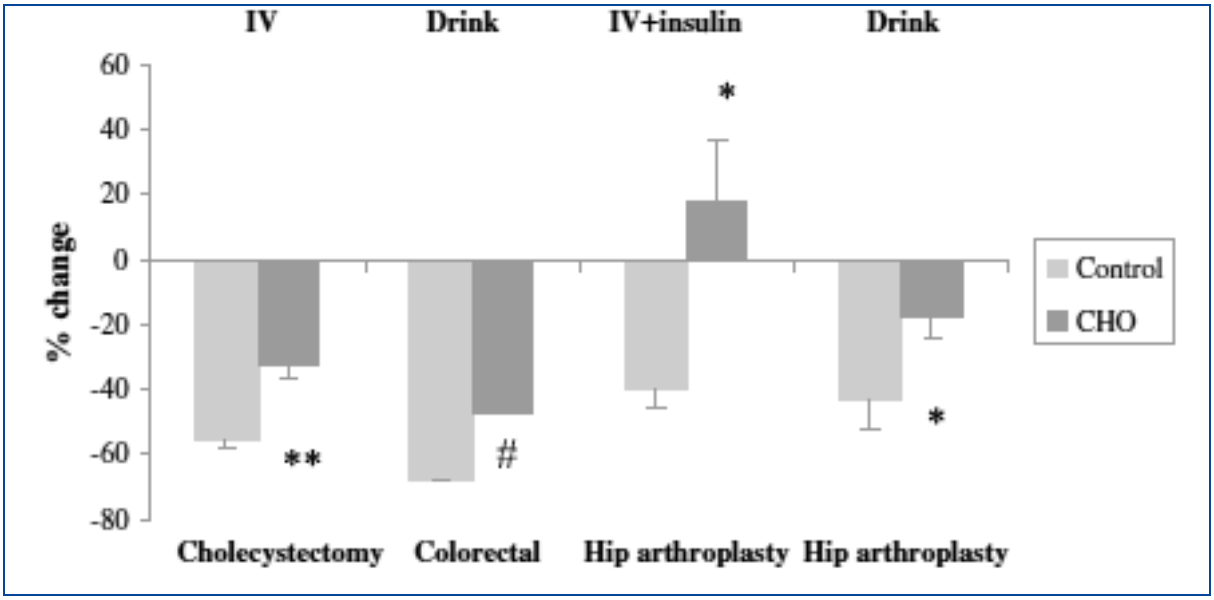


Figure 1. *The relative change in insulin sensitivity at postoperative measurement versus before surgery in patients fasted overnight before surgery (Control) and patients given carbohydrates (CHO) as an intravenous infusion (IV), an intravenous infusion together with insulin before and during surgery (IV + insulin), or as a carbohydrate-rich beverage 2 hr before surgery (Drink).*

J. Nygren. Metabolic effects of fasting and surgery. Best Practice & Research Clinical Anaesthesiology. Vol. 20, No. 3, pp. 429-438, 2006

The effects on postoperative glucose production by avoiding preoperative fasting indicate that there may also be effects on protein breakdown and gluconeogenesis. In support of this, other placebo-controlled, randomized studies have shown better maintained lean body mass (mid-arm circumference) and less reduced muscle strength (voluntary quadriceps strength) after gastrointestinal surgery when preoperative fasting was avoided. The avoidance of preoperative fasting has not been associated with differences in cytokine or stress hormone levels.

The metabolic effects of providing fluid and energy before surgery using a carbohydrate-rich drink versus placebo was also shown in randomized trials to affect patient well-being (thirst, hunger, anxiety) before surgery and to reduce postoperative nausea and vomiting (PONV) after laparoscopic cholecystectomy.

Indeed, there are also other important determinants of the surgical stress response. Both minimally invasive procedures as well as epidural anaesthesia have been shown to reduce postoperative insulin resistance. Mechanisms related to these effects may be a reduction in cytokine responses (laparoscopic surgery) and lower release of stress hormones (epidural anaesthesia). In studies where patients undergoing open colorectal surgery were treated with an enhanced recovery protocol, including avoidance of preoperative fasting and epidural anaesthesia, postoperative insulin sensitivity was well maintained.

Controversial patient groups

- Obesity

Harter et al in Anaesthesia & Analgesia (1998) compared 256 fasted patients (75 obese and 157 lean). They found that in the obese patients who were otherwise healthy, there was no increase incidence of increased gastric volumes and low stomach pH contents in the fasted state.

In addition, Maltby et al in the Canadian Journal of Anaesthesia (2004) studied 126 patients who were obese ASA 1 or 2 patients. They also could not demonstrate an increase in gastric volume or lower pH in these patients after being routinely fasted. They concluded that obesity per se should not be considered a risk factor for pulmonary aspiration and that healthy obese patients should be allowed to drink clear fluids up to 2 hours before elective surgery.

- Diabetes

Gustafsson et al in Acta Anaesth Scand (2008) studied gastric emptying time of 25 type 2 diabetics vs 10 healthy control subjects. They were given carbohydrate rich drink with paracetamol 1.5g. The diabetic patients demonstrated elevated glucose levels at the 2 hour mark but this had returned to normal at 3 hours. There was no difference in gastric-emptying half time. They concluded that the preoperative carbohydrate rich drinks could be used safely in diabetics but that these drinks should be given 3 hours preop instead of 2 hours.

Conclusion

ERAS protocols strongly recommend the use of carbohydrate-rich drinks in patients undergoing major elective surgery. Although this approach does not appear to cause any harm to patients (aspiration risk), its clinically relevant outcome benefit is supported by reasonably low quality evidence. There is however no doubt that unnecessarily prolonged starvation times causes patient discomfort and is potentially detrimental to postoperative patient outcomes.

Key Words

- **Glycogenolysis** - glycogen breakdown to glucose
- **Gluconeogenesis** - glucose synthesis from non-carbohydrate sources – lactate, pyruvate, amino acids & glycerol. Enhanced by glucocorticoids, catecholamines, glucagon and thyroid hormone, whereas insulin inhibits it.
- **Proteolysis** - the breakdown of proteins or peptides into amino acids by the action of enzymes
- **Glycolysis** - breaks down glucose and forms pyruvate with the production of two molecules of ATP. The pyruvate end product of glycolysis can be used in either anaerobic respiration (if no oxygen is available) or in aerobic respiration via the citric acid cycle which yields much more usable energy for the cell
- **Lipolysis** - the breakdown of lipids and involves hydrolysis of triglycerides into glycerol and free fatty acids.

References

1. C. Simons. Preoperative Fasting Guidelines in the Age of ERAS. UCT Part II Anaesthesia Refresher Course. 2015
2. J. Nygren. Metabolic effects of fasting and surgery. Best Practice & Research Clinical Anaesthesiology. Vol. 20, No. 3, pp. 429-438, 2006
3. P.C. Stuart The evidence base behind modern fasting guidelines. Best Practice & Research Clinical Anaesthesiology. Vol. 20, No. 3, pp. 457-469, 2006
4. E. Søreide, O. Ljungqvist. Modern preoperative fasting guidelines: a summary of the present recommendations and remaining questions. Best Practice & Research Clinical Anaesthesiology. Vol. 20, No. 3, pp. 483-491, 2006
5. S. Awad, K. Varadhan, O. Ljungqvist, D. N. Lobo. A meta-analysis of randomised controlled trials on preoperative oral carbohydrate treatment in elective surgery. Clinical Nutrition 32 (2013)
6. Gokhan Yagci et al. Effects of preoperative carbohydrate loading on glucose metabolism and gastric contents in patients undergoing moderate surgery: A randomized, controlled trial. Nutrition 24 (2008) 212–216
7. J. Roger Maltby. Fasting from midnight - the history behind the dogma. Best Practice & Research Clinical Anaesthesiology. Vol. 20, No. 3, pp. 363-378, 2006
8. M. J. Scot et al. Enhanced Recovery After Surgery (ERAS) for gastrointestinal surgery, part 1: pathophysiological considerations. Acta Anaesthesiologica Scandinavica 59 (2015) 1212–1231
9. A. Feldheise et al. Enhanced Recovery After Surgery (ERAS) for gastrointestinal surgery, part 2: consensus statement for anaesthesia practice. Acta Anaesthesiologica Scandinavica 60 (2016) 289–334
10. Olle Ljungqvist. Modulating postoperative insulin resistance by preoperative carbohydrate loading. Best Practice & Research Clinical Anaesthesiology 23 (2009) 401–409
11. Kotze V. Perioperative nutrition: what do we know? S Afr J Clin Nutr 24(3) Supplement (2011)
12. Smith MD et al. Preoperative carbohydrate treatment for enhancing recovery after elective surgery (Review). Cochrane Database of Systematic Reviews. 2014.

Central Venous Pressure and the Pulmonary Artery Catheter

Dr Frank Schneider

UCT Dept of Anaesthesia & Perioperative Medicine

Central venous pressure (CVP) and pulmonary artery pressure (PAP), used as an indication of cardiovascular filling, have been long-standing components of haemodynamic monitoring and optimisation. Both central venous catheters (CVC) and pulmonary artery catheters (PAC) require central vein access, with the inherent risks and complications associated with this invasive procedure. Additional complications related to transcardiac passage and the pulmonary vasculature are relevant to PAC use. The following notes will summarise the important aspects of CVP and PAP monitoring in routine clinical use.

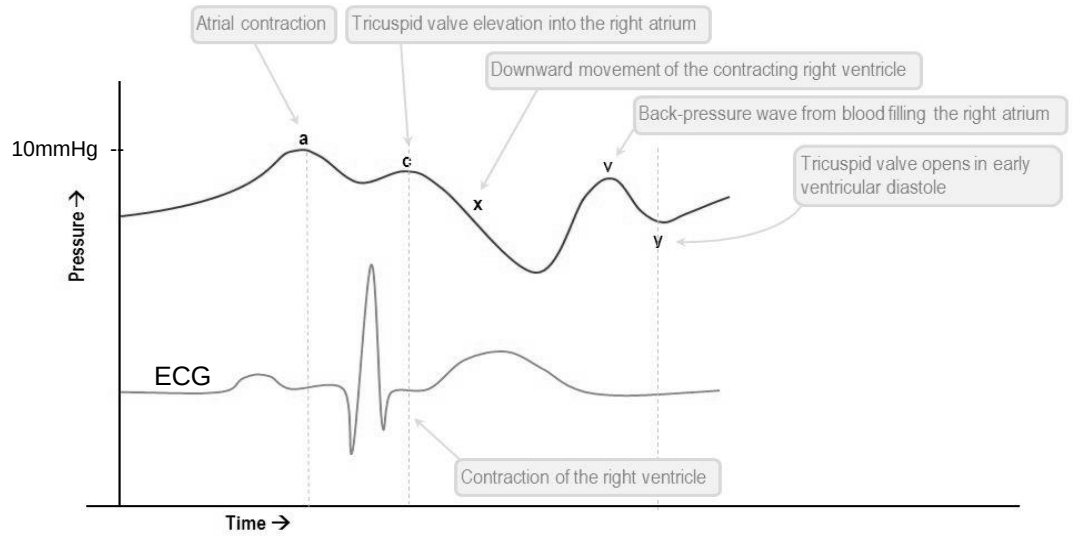
Central Venous Pressure

This is the pressure within the intrathoracic vena cava, is usually measured by the insertion of a CVC in the internal jugular or subclavian vein, and equates to right atrial pressure in the absence of caval obstruction. The point of reference and zeroing of a transducer is at the level of the right atrium, at the 4th intercostal space in the mid-axillary line ("phlebostatic axis" when supine). By convention, pressure should be recorded at the base of the c-wave, at end expiration in the supine position, which equates to right atrial pressure just prior to ventricular systole.

The normal CVP waveform is divided into:

> **'a' wave** – caused by increased right atrial pressure during atrial contraction and correlates with the P wave on the ECG. It disappears in atrial fibrillation or may be seen as flutter waves in atrial flutter. The descent of the 'a' wave occurs with atrial relaxation and is interrupted by the 'c' wave.
>
> **'c' wave** – closure of the cusps of the tricuspid valve occurs as right ventricular pressure rises during isovolumetric contraction of the right ventricle, at the end of the QRS complex. Right atrial pressure rises due to the cessation of blood flow across the tricuspid valve.
>
> **'x' descent** – as the right ventricle contracts and the tricuspid annulus moves towards the cardiac apex (TAPSE – tricuspid annular plane systolic excursion, cf. as seen on echo), the right atrium is stretched and right atrial pressure decreases. This occurs before the T wave on ECG.

Figure 1 - Right atrial pressure waveform (adapted from www.derangedphysiology.com)

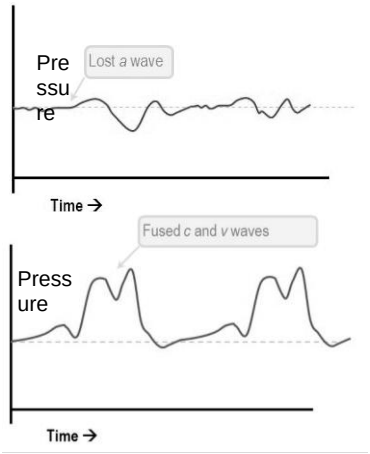


'v' wave – reflects rise in atrial pressure with continued blood inflow against a closed tricuspid valve throughout late systole and early diastole (after the T wave on ECG) and distension of the right atrium.

'y' descent – the tricuspid valve opens as soon as right atrial pressure exceeds right ventricular pressure, causing a rapid drop in pressure as atrial blood flows into the right ventricle, just before the P wave on the ECG. The third heart sound (S3) corresponds to the nadir of the 'y' descent, with the subsequent ascent reflective of continued atrial filling.

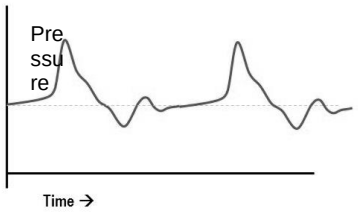
Abnormal CVP waveform patterns can be seen in the following circumstances:

Atrial fibrillation (AF) or flutter. As the 'a' wave is caused by atrial contraction, it disappears during atrial fibrillation, or may be appreciated as flutter waves in atrial flutter. Long-standing AF may lead to atrial dilatation and rise in right atrial pressure.

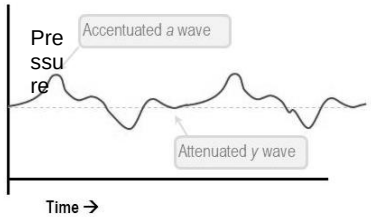


Tricuspid regurgitation. During right ventricular systole, blood is ejected through an incompetent tricuspid valve in to the right atrium, causing fusion of the 'c' and 'v' waves, and obliteration of the 'x' descent (as it reverses the effect of TAPSE). Mild regurgitation may not affect the waveform pattern.

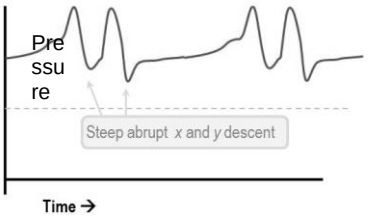
Atrioventricular dyssynchrony. Fusion of the 'a' and 'c' waves can occur during junctional rhythms, as atrial and ventricular contraction occur simultaneously. Cannon 'a' waves occur when atrial contraction coincides with a fully closed tricuspid valve, as seen during complete heart block, premature beats, or retrograde conduction of ventricular depolarization (e.g. ventricular tachycardia).



Reduced right atrial emptying. Increased resistance to right atrial emptying during atrial systole produces a large 'a' wave and attenuated 'y' wave. This may be due to flow obstruction at the tricuspid level (tricuspid stenosis, atrial myxoma or thrombus, carcinoid disease), or beyond (right ventricular hypertrophy, pulmonary stenosis, pulmonary hypertension).



Pericardial constriction. In constrictive pericarditis and restrictive cardiomyopathy the CVP is raised and a sharp 'y' descent, or bifid waveform with abrupt 'x' and 'y' descents may be seen. In cardiac tamponade, CVP is elevated, 'x' and 'y' descents are not prominent, and the 'y' descent is prolonged.



www.derangedphysiology.com

With a competent tricuspid valve, right atrial pressure (equal to CVP in the SVC) should equal right ventricular end-diastolic pressure and may predict preload when right ventricular compliance is normal. However, the variation in right ventricular compliance amongst many other factors such as intrinsic systolic and diastolic myocardial function, pulmonary vascular resistance, and positive pressure ventilation, has shown CVP to be a poor predictor of fluid responsiveness. At best, CVP should be used as a trend value and correlated with the individual clinical context.

Pulmonary Artery Catheter

The Swan-Ganz catheter (or PAC), was introduced by Swan and Ganz in 1970 when they reported catheterization of the heart using their flow-directed balloon-tipped catheters. Their use in intensive care has decreased significantly in recent years, following multiple publications of lack of evidence for benefit or harm in critically ill patients, notably the PAC-Man trial in 2005. The development of less invasive cardiac output monitors has also contributed to the decline in PAC use, as has the marked increase in utilisation of echocardiography, which can provide most of the haemodynamic data available from the PAC. However, in complex patients with known pulmonary hypertension, severe left or right ventricular dysfunction, or severe valvular disease, the PAC may still be justified with echocardiography as a complementary technique.

Suggested indications for PAC placement might include:

- Investigation and quantification of cardiac shunts
- Perioperative monitoring: complex coronary and valvular surgery, vascular surgery, solid organ transplantation
- Critical care: severe left or right ventricular failure, pulmonary hypertension, constrictive pericarditis, valvular heart disease, obtaining mixed venous oxygen saturations, continuous cardiac output monitoring
- Therapeutic: drug delivery to pulmonary circulation (e.g. prostacyclin), cardiac pacing

Contraindications to insertion include prosthetic tricuspid or pulmonic valves, right-sided valvular vegetations or thrombi and intracardiac right mass.

Enthusiast centres continue to use PACs routinely, with good results and minimal complications. Furthermore, PACs are still the gold standard against which novel cardiac output monitors are compared. Emphasis should be put on using haemodynamic data, irrespective of how they are obtained, to inform therapeutic decision-making and appropriate patient management.

Catheter
Modern adult PACs are 7 or 7.5 French gauge (Fr) and 110cm long, with 10 cm markings along the length to aid insertion, four or five lumens, and an inflatable 1.5ml balloon at the tip of the catheter. A blue lumen and proximal injectate port terminates at 30cm from the tip of the catheter, and should lie within the right atrium when the PAC is correctly sited. The white proximal infusion port terminates at 31cm from the tip, should also lie in the right atrium, and can be used for fluids and medication. The yellow distal port is the pulmonary artery lumen and allows measurement of pulmonary artery (PA) pressures and mixed venous saturation. A shorter red lumen connects to the balloon and allows inflation or deflation, with a locking mechanism to prevent deflation during flotation. A thermistor is located 4cm from the catheter tip, is attached to a proximal thermistor connector and used for thermodilution cardiac output measurement. Additional connectors may allow pacing capabilities, continuous cardiac output and oximetry measurement via fibreoptic sensors.

Insertion
The catheter is inserted using a sterile Seldinger technique similar to CVC insertion, within a sterile cover (to allow manipulation of the PAC) via an 8 or 8.5Fr introducer sheath with a high volume sideport. The PA transducer is attached to the distal lumen to display waveform changes during flotation.

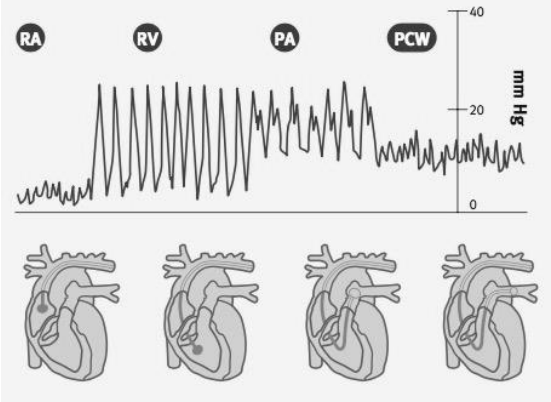


Figure 2 – Pressure waveform during advancement of PAC

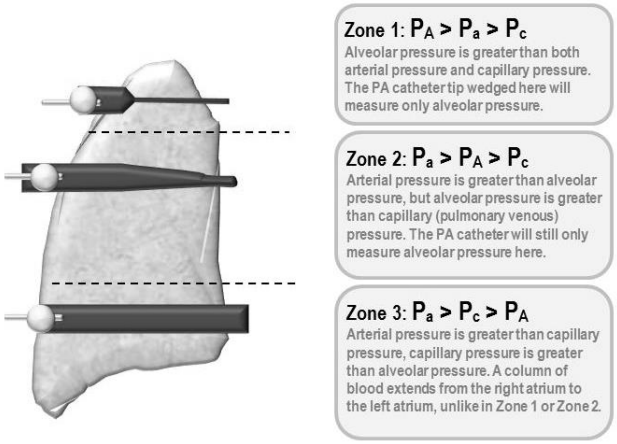
The catheter, in adults, is advanced 15-20cm from the jugular vein (10-15cm from subclavian, 30-40cm from femoral) to the right atrium. Following balloon inflation (usually max 1.5ml air), the catheter is advanced a further ~10cm, whilst watching for ectopic beats and a change to the right ventricular waveform. Systolic pressure is higher in the right ventricle (15-30mmHg) with diastolic pressure similar to right atrial pressure (1-6 mmHg).

The catheter should now be advanced a further ~10cm in order to float into the pulmonary artery. Pulmonary artery systolic pressure should be almost equivalent to right ventricular systolic pressure. Closure of the pulmonic valve causes a dicrotic notch on the pulmonary artery tracing, and special attention should be paid to the diastolic pressures, which are higher in the pulmonary artery. The catheter is advanced further until it is "wedged" in a central branch of the pulmonary artery, and now measures pulmonary artery wedge pressure (PAWP) or pulmonary artery occlusion pressure (PAOP), and the waveform shown will be that of left atrial pressures.

Insertion of the PAC may be hampered by tricuspid regurgitation or coiling in the right ventricle (RV). Transoesophageal echocardiography (TOE) allows real-time visualisation of the passage of the PAC from the right atrium, through the RV and into the PA (from a mid-oesophageal RV inflow-outflow view), and confirmation of the catheter tip in the pulmonary artery (>90% float into the right PA). Depending on the clinical setting, fluoroscopy or chest x-ray may also be used to confirm placement, as well as location within West Zone IIII.

Pulmonary artery wedge pressure

This represents left atrial filling pressure (normally 6-12 mmHg) and in patients with normal physiology estimates left ventricular end-diastolic pressure (LVEDP), which gives an indication of left ventricular end-diastolic volume (LVEDV), but – like CVP – does not estimate preload or predict cardiac performance or fluid responsiveness. The catheter tip should be in West Zone III for accurate measurements, ensuring a static column of blood between the catheter tip in the PA and the left atrium. PAWP will be greater than LVEDP in mitral stenosis or regurgitation, with atrial myxoma, pulmonary venous obstruction (due to fibrosis or vasculitis), left-to-right shunt, COPD and positive pressure ventilation. PAWP will be less than LVEDP in left ventricular failure, raised intra-thoracic pressure, decreased left ventricular compliance and aortic regurgitation.



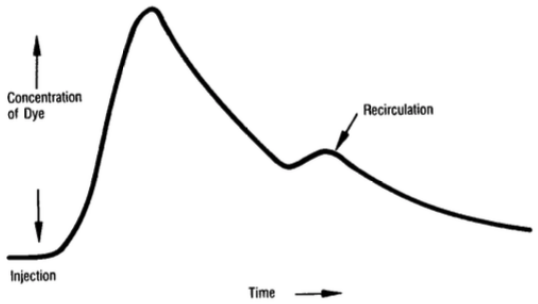
In addition to PAWP, the most common measured and derived parameters, with normal values, are as follows:

MEASURED VALUES	
Cardiac Output (CO)	4 – 8 l/min
Cardiac Index (CI)	2.5 – 4 l/min
CVP	2 – 6 mmHg
PAWP	6 – 12 mmHg
PA Pressure (PAP)	25/10 mm Hg
SvO_2 (mixed venous)	65 – 70%
Temperature	
DERIVED VALUES	
Stroke Volume (SV)	50 – 100 ml/beat
SV Index (SVI)	25 – 45 ml/beat/m^2
Systemic Vascular Resistance (SVR)	900 – 1300 dynes-sec/cm^5
SVR Index (SVRI)	1900 – 2400 dyne-sec/cm^5
Pulmonary Vascular Resistance (PVR)	40 – 150 dyne-sec/cm^5
PVR Index (PVRI)	120 – 200 dynes-sec.cm^5

Figure 3 Measured and derived values obtained with PAC

Cardiac output measurement

The indicator dilution method proposed by Stewart in 1890 for determining cardiac output, and later refined by Hamilton, relies on the upstream injection of a known amount of a substance and the measurement of the concentration of the substance over time by a downstream detector, giving an indicator dilution curve (shown alongside).

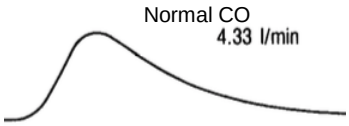


This was historically done using indocyanine green as the indicator, and cardiac output can be calculated using the Stewart-Hamilton equation:

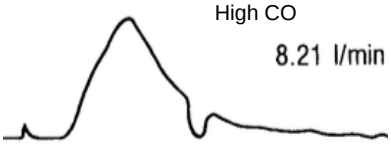
$$CO = \frac{I \times 60}{c_m \times t} \times \frac{1}{k}$$

where CO = cardiac output in l/min, I = amount of dye injected (mg), 60 = 60 sec /min, c_m = mean indicator concentration (mg/l), t = total curve duration (s) and K = calibration factor.

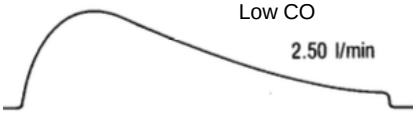
Thermodilution uses the same indicator dilution principles, using temperature change as the indicator, and has become the standard of practice for practical, bedside cardiac output measurement. Cold or room temperature solution is injected rapidly and the temperature change measured at the thermistor bead in the distal PAC. The change in temperature is negative, but is represented upright on a time-temperature curve by convention, where the area under the curve is inversely proportional to the cardiac output. A normal curve shows a sharp upstroke from rapid injection of cold fluid bolus, followed by a smooth curve returning to baseline.



With increased cardiac output, increased blood flow produces a steeper and shorter thermodilution curve.



Low cardiac output states will have a slurred curve with delayed peak, and slower return to baseline. This curve is exaggerated in tricuspid regurgitation with recirculation of injectate.



The simplified and modified Stewart-Hamilton equation becomes:

$$CO = \frac{V \times (T_b - T_i) K1 \times K2}{A}$$

where CO = cardiac output in l/min, V = injected volume, T_b = blood temperature, T_i = injectate temperature, K1 and K2 are corrections for specific heat and gravity of injectate and blood, and A = area under the curve (change in blood temperature as a function of time).

In order to ensure consistency and accuracy, injectate should be bolused rapidly at the end of expiration and averaged over a minimum of three measurements. Very cold injectate (0-4ºC) is more accurate (better signal to noise ratio), but may cause bradycardia and a decrease in cardiac output, making room temperature injectate safer.

Mixed venous oxygen saturation

Mixed venous oxygen saturation (S_vO_2) measurement is performed on a sample from the distal lumen of the PAC and is distinct from central venous oxygen saturation ($S_{cv}O_2$), which is usually from blood sampled in the superior vena cava (SVC). S_vO_2 gives an indication of oxygen delivery and extraction, but can be high (sepsis, hepatic failure, wedged PAC, high FiO_2, decreased oxygen demand) or low (multiorgan failure, anaemia, increased oxygen demand, cardiac arrest) in varied clinical contexts, making routine monitoring of limited clinical use.

References

1. Kumar A, Anel R, Bunnell E, et al. Pulmonary artery occlusion pressure and central venous pressure fail to predict ventricular filing volume, cardiac performance, or the response to volume infusion in normal subjects *Crit Care Med* 2004;**32**(3):691-699
2. Marik P, Cavallazzi R. Does the Central Venous Pressure Predict Fluid Responsiveness? An Updated Meta-Analysis and a Plea for Some Common Sense *Crit Care Med* 2013;**41**(7): 1774-81
3. Swan H, Ganz W, Forrester J, et al. Catheterization of the Heart in Man with Use of a Flow-Directed Balloon-Tipped Catheter *N Engl J Med* 1970; **283**:447-451
4. Sandham JD, Hull RD, Brant RF, et al. A randomized, controlled trial of the use of pulmonary-artery catheters in high-risk surgical patients. *N Engl J Med* 2003; **348:** 5–14
5. Harvey S, Harrison DA, Singer M, et al. Assessment of the clinical effectiveness of pulmonary artery catheters in management of patients in intensive care (PAc-Man): a randomised controlled trial. *Lancet* 2005; **366:** 472–7
6. Easby D, Dalrymple P. Monitoring arterial, central venous and pulmonary capillary wedge pressure. *Anaesthesia and Intensive Care Medicine* 2009; **10**(1):38-44

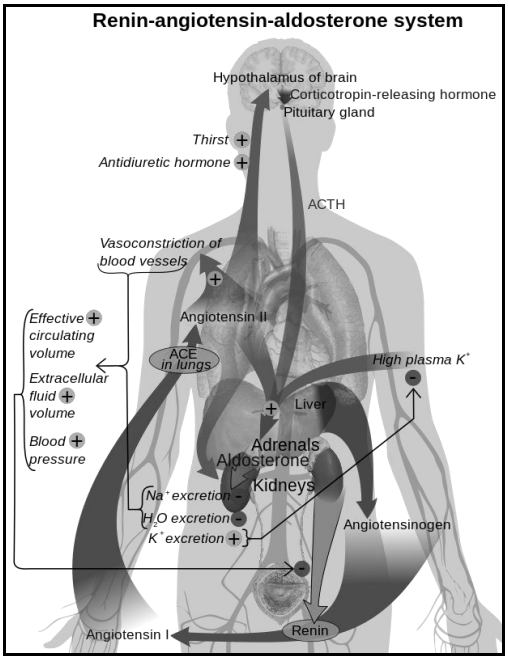
The ABC of Natriuretic Peptides

Dr Estie Cloete

UCT Dept of Anaesthesia & Perioperative Medicine

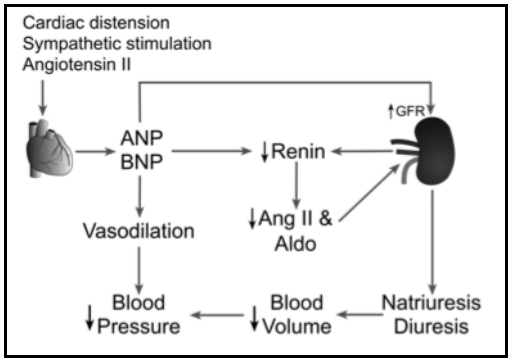
The natriuretic peptide system forms part of the regulation of extracellular fluid composition and volume. It not only impacts on the salt and water handling but also influence pressure regulation. Natriuretic peptides (NP) also have an influence on the myocardial structure and function. The main physiological action is a reduction in blood volume and systemic vascular resistance that lead to reduction in arterial pressure.

To understand the importance of the natriuretic peptides, we need to revise the RAAS (Renin-Angiotensin-Aldosterone-System).



The extracellular fluid volume is determined by the osmotically active solute in the extracellular fluid (ECF). Sodium and chloride are the most abundant solutes and the changes in chloride are usually secondary to changes in sodium. Thus the most important determinant of the ECF volume is sodium. The control of water excretion is controlled by volume, and a rise in the volume will inhibit Vasopressin secretion and vice versa. The osmotic regulation of Vasopressin secretion is overridden by volume stimuli.

Angiotensin II (ATII) has got an important role in the response to hypovolemia. Angiotensin II will stimulate Vasopressin and Aldosterone secretion. It causes blood vessel constriction and thirst that will help to maintain the blood pressure. With the expansion of the ECF volume, the natriuretic peptides ANP and BNP are secreted. This will lead to natriuresis and diuresis.



The focus will be on two natriuretic peptides secreted by the heart: ANP and BNP.

ANP

This was the first hormone isolated and is primarily released from atrial myocardial cells and in some cases the ventricles in response to atrial stretch from volume expansion. ATII, endothelin and sympathetic stimulation especially ß adrenergic mediated, carotid and baroreceptors also stimulates ANP release. Both atria participates in the release of ANP but it appears that the right is more important in quantity. ANP can be isolated from other tissues like the brain where ANP- containing neural pathway projects from the anteromedial part of the hypothalamus to the lower brainstem which controls neural regulations of the cardiovascular system. ANP opposes ATII and the ANP-containing neural circuits is involved in lowering the BP and promoting natriuresis.

BNP

Brain natriuretic peptide is homologous to ANP and initially identified in the brain but also present in the cardiac ventricles. This 32 amino acid molecule is present in circulation in concentrations 20% lower than that of ANP in normal subjects but exceed ANP in diseased states. The prohormone proBNP gets cleaved and then produces the active hormones BNP and N-terminal proBNP(NT-proBNP).

CNP

C-type natriuretic peptide is a paracrine mediator and is produced by vascular endothelial cells and present in the pituitary and kidneys with very little in the circulation and heart. The function may involve regulation in blood flow but its physiological role remains to be determined.

Mechanism of action

The two major actions of NP is

- Vasodilatory effects
- Natriuresis and Diuresis via the renal system

1. Vasodilator effects

ANP and BNP causes systemic vasodilation via cGMP on vascular smooth muscle. Venodilation with the increased venous compliance will lead to a decrease the central venous pressure and preload and thus a reduction in the cardiac output. Arterial dilatation leads to decrease in systemic vascular resistance (SVR) and arterial blood pressure

2. Renal effects

ANP and BNP increase sodium excretion by the kidney through dilatation of the afferent arterioles and relaxation of the mesangial cells. This then leads to an increase in glomerular filtration and inhibition of sodium reabsorption in the renal tubules. Additional actions of the natriuretic peptides include an increase in capillary permeability that leads to the extravasation of fluid and a decrease in blood pressure. In the RAAS these peptides inhibit renin secretion and counteract the pressor effects of the catecholamines and ATII.

Natriuretic peptide Receptors (NPR)

Three receptors have been identified: NPR-A, NPR-B and NPR-C. NPR-A and B is expressed on cell membranes and has guanyl cyclase domains. ANP has greatest affinity for NPR-A and CNP for NPR-B. NPR-C binds all 3 natriuretic peptides. It acts via G proteins, activates phospholipase C and inhibits adenylcyclase. Some sources however say that the receptors has no intracellular change and is a clearance receptor that removes natriuretic peptides form the circulation and releases them later to maintain a steady level of these hormones in the blood.

Summary of Cardiovascular and Renal Actions of Natriuretic peptides

- ❑ Arterial hypotension
- ❑ ↓ venous pressure
- ❑ Systemic vasodilation
- ❑ Natriuresis
- ❑ Diuresis
- ❑ ↑ GFR and filtration fraction
- ❑ Inhibition of renin release
    - ❑ ↓ ATII
    - ❑ ↓ Aldosterone
- ❑ ↓ Pulmonary capillary wedge pressure

To summarise: Natriuretic peptides is a counter-regulatory system for the RAAS.

Biomarkers

BNP as biomarker has important role in diagnosis, disease severity, risk stratification, therapeutic decision-making and prognosticator in the preoperative as well as postoperative period.

BNP binds to NPR-C and through proteolysis by the NEP, is cleared from the plasma. A small proportion of BNP and NT-proBNP is cleared by the kidneys. The half-life of BNP is 20 min and NT-proBNP, 120 minutes. Due to its wider detection range and its stability for 72 hours in the plasma, NT-proBNP is a more useful biomarker.

Their baseline values are influenced by a number of factors: higher values in female, elderly, anaemia and patients with atrial fibrillation and lower values in obese patients.

Clinical application in cardiac disease

ANP and BNP release is increased in cardiac failure (CF) in response to ventricular filling pressures. Both hormones have increased concentrations in patients with symptomatic LV dysfunction as well as asymptomatic patients. ANP and BNP have diuretic, natriuretic and hypotensive effects as well as inhibit the RAAS system, endothelin secretion and renal sympathetic activity. ANP and BNP in patients with heart failure will counteract the effects of noradrenaline, endothelin and ATII that will limit the vasoconstriction and sodium retention. BNP has an added role in protecting the heart against collagen accumulation and pathological remodelling in progressive cardiac failure.

The values of BNP and NT-proBNP in CF correlate with disease severity and functional class of New York Heart Association. It is inversely related to the cardiac output.

NT-pro-BNP is a prognosticator of unfavourable outcome such as cardiovascular death, readmission and cardiac events in chronic heart failure patients as well as those with asymptomatic left ventricular dysfunction.

This table[3] demonstrates the cut off values for diagnostic purposes in cardiac failure (CF).

	Rule out (CF unlikely)	Rule in (CF likely)
BNP (pg/ml)	100	500
NT-proBNP (pg/ml) age < 50 yr	300	450
NT-proBNP (pg/ml) age > 50 yr	300	900

In patients with ischaemic heart disease the degree of myocardial damage is linked to the rise in NT-proBNP and this correlates with the LV ejection fraction as well.

As marker in maternal cardiac disease, NT- proBNP is an early marker of decompensation and worsening of cardiac disease. Patients with pre-eclampsia has a significantly higher level than the normotensive patients due to elevated LV filling pressures and the underlying LV diastolic dysfunction.

The ACC/AHA HF guidelines recommend the measurement of natriuretic peptide levels in evaluation and risk stratification in the urgent care setting in whom clinical diagnosis of CF is uncertain. It is part of the total evaluation but should not be used in isolation to confirm/exclude suspected CF. The use of BNP in differentiating cardiac vs. non-cardiac cause of dyspnea is difficult and remains uncertain. It may be useful as part of clinical evaluation in patients with dyspnoea of uncertain etiology and a history of cardiomyopathy induced CF versus a non-cardiac cause of dyspnoea.

Perioperative use

In cardiac surgery especially in the patient with aortic stenosis, the levels of BNP and NT-proBNP correlate with severity but also decline after successful valve replacement. Pre-operative levels correlate with the long-term outcome after cardiac surgery.

NT-proBNP in the non-cardiac surgery setting, predict mortality, as well as the Major Adverse Cardiac Events (MACE) peri-operatively. Perioperative myocardial ischemia is difficult to diagnose, and its early recognition, as well as the detection of heart failure versus dyspnoea from a pulmonary origin, might be guided by post-operative NT-proBNP levels. This is important to make early therapeutic decisions in these patients.

There is good evidence to support the perioperative use of BNP as biomarker in diagnosis, risk assessment and prognosticator. Ongoing research will assess whether the perioperative use NT-proBNP levels will influence patient outcome.

Pharmacology

Neutral endopeptidase (NEP) degrades natriuretic peptides and inhibition of this enzyme increases the circulating levels of natriuretic peptides and potentiate their effects. In animal studies NEP inhibitors is effective for the treatment of heart failure especially when combined with an ACE inhibitor.

Recombinant BNP, Nesiritide, is used for the acute decompensated congestive cardiac failure due to systolic dysfunction.

To summarise: Natriuretic peptides have an important role in counter regulation of the RAAS system and can be used as biomarkers with a wide application in clinical use.

References

1. Barrett KE, Barman SE, Boitano S, Brooks HL. Ganong's Review of Medical Physiology 25e (2016) McGraw-Hill Education.
2. Klabunde RE. Cardiovascular Physiology Concepts: Atrial and Brain Natriuretic peptides. http://www.cvphysiology.com/Blood%20Pressure/BP017
3. Afshani N, BM Biccard, Dyer RA. (2012). The role of biomarkers and B-type natriuretic peptide in diagnosis and perioperative risk prediction. Southern African Journal of Anaesthesia and Analgesia. 18:6, 303-308.
4. https://en.m.wikipedia.org/wiki/Renin–angiotensin_system#/media/File%3ARenin_angiotensin_aldosterone_system.png

Blood Products

Dr Rebecca Gray

Division of Paediatric Anaesthesia
UCT Dept of Anaesthesia & Perioperative Medicine

The South African National Blood Transfusion Service (SANBS) issues over a million blood components per year. This brief overview of blood products will touch on screening, testing, processing, storage and clinical use of the major products.

Whole blood is collected from a donor into a bag prepared with sodium citrate, a stable, minimally toxic anticoagulant with pH buffering properties. It is chilled and transported to the blood processing facility where it is tested and processed into the various component products and thereafter stored under conditions specific to each product until required for clinical use.

Operating in a region with one of the highest HIV infection rates in the world has made accurate screening of donor units for infection a priority. Stringent donor selection procedures together with a policy of non-remuneration for donations and the use of volunteer donors only are the first steps in the process. Following collection blood is screened using advanced nucleic acid and serologic testing and the residual risk of transmitting HIV, HCV and HBV infection is remote.

After screening whole blood is centrifuged to separate the plasma, buffy coat (leucocytes and platelets) and red cells. The red cells are removed at this point to prepare red cell concentrate products. Removing the buffy coat and plasma removes 70-80% of the leucocytes from the red cell product, reducing the risk of non-haemolytic febrile transfusion reactions.

RED CELL PRODUCTS

Table 1: Packed Red Cells

What's in the bag	Sodium Citrate (anticoagulant)		
	111ml preservative solution containing:		
	Adenine	helps maintain ATP levels during storage	
	Glucose	provides a substrate for RBC energy pathways	
	Saline and *mannitol*	reduce the haemolysis of the banked red cells during the 42-day storage period	
Shelf life	@ 1-6°C up to 42 days		
Storage lesion		Day 1	Day 35
	pH	7.5	6.7
	K^{+*}	4.0 mmol/L	66 mmol/L
	Na^{+}	155 mmol/L	122 mmol/L
	Dextrose	24.5 mmol/L	5 mmol/L
	NH_3	100 mg	900 mg
	2,3 DPG*	13 mmol/L	<1 mmol/L
	Red cell survival	100 %	70 %
Compatibility testing	- ABO matching - Rh- patients should get Rh- blood as far as possible. - Antigen negative blood should always be transfused to patients with specific and clinically significant red cell antibodies.		
Dose	4ml/kg to raise Hb by 1g%		
	Packed cells (300±50ml)		R1217
	Packed cells (Leucocyte depleted) (260±50ml)		R2115.79
	Packed cells washed		R4028

*see notes below

*At standard storage temperatures of 1-6ºC, the sodium-potassium pump is essentially non-functional and intracellular and extracellular levels gradually equilibrate. Plasma **potassium** concentration increases nearly eightfold over 28 days of storage although, at expiry, the total potassium load in red cell concentrates is only about 9.5 mmols.

*Red cell **2,3 DPG** decreases with storage, levels drop significantly by 1 week, shifting the oxygen-dissociation curve to the left and decreasing tissue oxygen delivery. Full restoration of 2,3,DPG takes about 72 hours but there is 50% recovery within 7 hours of transfusion. In clinical situations of hypoxia and lactic acid production, and with decreasing pH, the oxygen dissociation curve is shifted to the right, increasing oxygen delivery. Increased oxygen delivery also occurs with an increase in cardiac output. It is therefore generally considered that low 2,3 DPG levels in stored blood are not usually clinically significant HOWEVER patients in shock who cannot increase cardiac output to compensate, patients receiving large volumes of stored blood such as occurs in massive transfusion, or in patients undergoing red cell exchange procedures, transfusion of blood which has been stored for less than 5 days may be optimal.

Transfusion related metabolic or electrolyte derangements

-*Citrate toxicity**:*** Following transfusion the citrate from the unit is metabolized in the Krebs cycle of respiration in most cells in the body, particularly in the liver, muscle and renal cortex. In certain clinical conditions (liver disease, hypothermia and hypoparathyroidism; neonates with inadequate calcium stores and/or immature livers) patients be at increased risk for 'citrate toxicity' during rapid transfusion of red cells or FFP. In such cases citrate may precipitate arrhythmias and reduced cardiac contractility by binding plasma calcium thereby lowering plasma ionized calcium levels. With rapid transfusion or in clinical cases at risk of citrate toxicity, ionised calcium levels should be monitored and 10 ml of 10% calcium gluconate administered intravenously: 10 ml for every 2 units of whole blood given in under 10 minutes or 0.5ml/kg for paediatric patients with proven (or suspected) hypocalcaemia.

- *Hyperglycaemia* due to glucose containing solution (especially in cardiac surgery in infants, transfusions during liver transplant). Problem is greater with fresher cells.

- *Hyperkalaemia*

Washed Red Cells

These are prepared on demand from banked packed cells. Cells are suspended in isotonic saline, centrifuged, saline removed and then re-suspended in isotonic saline for transfusion. They should be administered within 24 hrs because of the increased likelihood of bacterial contamination in preparation and the lack of nutrient admixture. Washed cells are indicated in the following circumstances:

1. severe, recurrent allergic transfusion reactions
2. known IgA deficiency in patients who have formed anti-IgA antibodies
3. Stored red cells which have been gamma irradiated (washes off the potassium which collects following irradiation, best managed by transfusing as soon as possible after irradiation)
4. paroxysmal nocturnal haemoglobinuria
5. Neonates with T-activated red cells (red cell T-crypt antigens have been exposed by bacterial infection particularly associated with necrotizing enterocolitis)

Evidence is mounting that the last two indications may no longer be relevant. In the case of neonates with T-cell activation, increasing use of packed cells (which have minimal plasma) as opposed to whole blood means it is probably unnecessary to provide washed red cells as a routine.

Frozen Red Cells

These are used in the case of unusual phenotypes (for patients with rare red cell phenotypes or multiple red cell antibodies) or in the case of autologous donation when donations are made over an extended period (beyond the shelf life of the liquid preserved red cells). This modification is prepared by adding glycerol, a cryoprotective agent, to red cells before freezing. Frozen red cells may be stored for up to 10 years and for longer intervals if there is a particular need for specific units. The thawed unit should be deglycerolized prior to transfusion. This is done by washing the cells with sodium chloride. The washed red cells are then re-suspended in additive solution and have a 24 hour shelf life.

This facility is not offered at the WPBTS or SANBS.

Whole Blood

When blood is harvested it is separated into component therapy soon after arrival in the processing plant. If it is kept as whole blood there is rapid degradation of the platelets and the clotting factors rendering them ineffectual within hours to days of donation. Whole blood is stored as for packed cells and has a shelf life of approximately 28 days. Whole blood may be considered in cases of exchange transfusion or massive haemorrhage but in the latter case, where coagulation may be a concern, component therapy may still be preferable.

PLATELETS

Table 2: Platelets

Preparation	- Derived from the buffy coat layer, separated within 8 hours of donation. - May be irradiated with no loss of function.	
	Pooled: 4-5 donations are pooled, re-suspended in plasma or platelet additive solution (PAS): 200-300ml containing ≥2.4x10^{11} plt	
	Single donor apheresis concentrate: from a single donor, up to 3 bags collected per donation. Automatically leuco-depleted.	
Storage	Stored at 20 - 24º on a platelet agitator	
Shelf Life	5 days	
Compatibility Testing	Ideally ABO compatible. In cases of severe shortages, can use non-identical (ideally group O), this may result in reduced platelet increase. RhD –ve should be given to premenopausal RhD -ve women.	
Indications	Thrombocytopaenia associated with actual or potential bleeding	
Dose	10ml/kg should raise the platelet count by 50x10^{9}	
Cost	Pooled platelet	R8507
	Leucodepleted	R10355
	Single donor mega unit	R11418
	Infant platelet (50ml)	R2004.39

Platelet Refractoriness

Repeated failure to obtain satisfactory elevation of platelet count despite multiple transfusions. HLA allo-immunisation is the main immune related cause, (leukocyte depletion lowers the risk). Non-immune causes include fever, infection, drugs, splenomegaly and DIC. Patients who are likely to require repeated platelet transfusions (e.g. ITP) should receive leucocyte depleted concentrates and be exposed to as few donors as possible – this is best achieved by using single donor apheresis concentrates. HLA matched concentrates are a possible solution when allo-immunisation occurs.

Adverse Effects with Platelets

Febrile reactions. Risk of bacterial contamination is greater because platelets stored at room temperature.

PLASMA PRODUCTS

Plasma components are derived by physical separation methods (e.g. centrifugation) and include FFP and cryoprecipitate whereas plasma derivatives are derived from large pools of plasma by more complex physical and chemical processes (e.g. alcohol based fractionation). Products derived from a pool of 12+ donations are classified as medicines and need to be registered. All plasma products are potentially antigenic.

Fresh Frozen Plasma (FFP)

Table 3: FFP

Preparation	Separated by centrifugation within 18 hrs of collection	
What's in the bag*	Glucose	24.8 mmol/L
	Potassium	3.2 mmol/L
	Sodium	165 mmol/L
	Chloride	79 mmol/L
	Osmolarity	322 mmol/L
	pH	7.9
Storage	- -18^{0}C (thus expect normal physiological levels of coagulation factors) -Thaw at 30-37^{0}C then transfuse within 4 hours (labile factors deteriorate within hours)	
Crossmatch	-ABO matched or blood with low anti-A or anti-B titres -O FFP should only be given to O-group patients -RhD sensitisation is unlikely	
Indications	-factor deficiencies where specific factor concentrate is not available -Multiple factor deficiencies with active bleeding (DIC, liver disease) and coagulopathy confirmed with point-of-care testing -Warfarin toxicity -Vitamin K deficiency with active bleeding -Suxamethonium apnoea	
Dose	10 – 15ml/kg,	
Cost	FFP (280±70ml)	R1193
	Leucocyte reduced FFP ((280±70ml)	R2404

* FFP is hyperosmolar, hypernatraemic and hypokalaemic and may precipitate pulmonary oedema and electrolyte imbalances if large volumes are transfused.

Cryoprecipitate

Table 4: Cryoprecipitate

Preparation	Extracted when FFP is thawed to 0 – 4 ^{0}C	
Storage	-18^{0}C for up to 1 year	
Active factors	- Factor VIII/vWF - Fibrinogen - Fibronectin - Factor XIII	
Cross matching	Not required	
Indications	- Congenital or acquired hypofibrinogenaemia (e.g. massive haemorrhage with fibrinogen <1.5g/L) - Factor XIII deficiency	
Dose	1 unit/ 5 - 10kg	
Cost	Cryo 100iU (10±1 ml)	R709

Cryosupernatant a.k.a. cryo-poor FFP is the component available following extraction of cryoprecipitate from FFP. The only indication is for use in therapeutic plasma exchange in the management of TTP.

PLASMA DERIVATIVES

Freeze-Dried Plasma (Bioplasma® National Bioproducts Institute) is produced from pooled virally inactivated fresh plasma using cold ethanol fractionation to separate the plasma into intermediate protein fractions. The various fractions undergo purification by precipitation, centrifugation and filtration. Formulations and sterile filling processes are employed to produce sterile final products. It should be stored at room temperature and has a shelf life of several years making it useful for pre-hospital conditions. It is reconstituted with sterile water and should be infused immediately through a blood giving set in a line with no residual calcium containing fluid. Rapid infusions may precipitate citrate toxicity.

100ml of FDP contains 4 – 6g plasma proteins with a normal distribution of human plasma components including albumin, immunoglobulins, coagulation and complement factors and their inhibitors. Coagulation factor activity is reduced by approximately 15% compared to FFP in animal models.

Albumin is available in 4% and 20% solutions. Prepared from pooled plasma using ethanol fractionation, sterilised by filtration and pasteurised by heat for 10 hours at 60°C to inactivate any HIV, HBV and HCV that may not have been detected in the initial screening of the donations. 20% solutions are indicated for hypoproteinaemia where as 4% solutions are used in certain contexts for plasma volume expansion. Care should be exercised where large volumes are administered as hyponatraemia may ensue:

Table 5: Albumin solutions available in SA

PRODUCT	VOLUME	ALBUMIN CONTENT	STABILISERS	[Na^+]	[K+]	[Citrate]	pH
Albusol 4%	200ml	8g/200ml	Sodium carpylate, 3 % dextrose	130	2	<4	7.0
Albusol 20%	50 or 100ml	20g/100ml	acetyl tryptophanate, sodium caprylate	<100	<10	<20	7.0
WPBTS 20% albumin	50 or 100ml	20g/100ml	Sodium carpylate	<130	<19		7.0

COAGULATION FACTOR CONCENTRATES

Table 6: Coagulation factor concentrates available in SA

PRODUCTS	CONTENT	UNITS	INDICATION
VIAHF 250 (*WPBTS*)(Paediatric)	Factor VIII/vWF	250 IU FVIII	Haemophilia A Von Willebrand disease
VIAHF 500 (*WPBTS)*(Adult)	Factor VIII/vWF	400-600 IU FVIII	Haemophilia A Von Willebrand disease
Haemosolvate Factor VIII 300 IU (*NBI*)	Factor VIII/vWF	300 IU FVIII/vWF	Haemophilia A Von Willebrand disease
Haemosolvate Factor VIII 500 IU (*NBI*) Two pack sizes: 500 IU and 2x 500 IU	Factor VIII/vWF	500 IU FVIII/vWF	Haemophilia A Von Willebrand disease
Haemosolvex Factor IX	Factor IX, Factor II, Factor VII, Factor X	500 IU FIX	Haemophilia B

LEUCOCYTE DEPLETION (LD):

Leucocytes distinguish self from foreign cells on the basis of Human Leucocyte Antigen (HLA) protein on the cell membrane. There is an enormous diversity of HLA proteins expressed on individual leucocytes and the chance of two unrelated individuals with identical HLA molecules on all loci is very low. With allogenic blood transfusion there will be tremendous exposure to these antigens and this is responsible for many of the non-haemolytic transfusion reactions seen. The leucocyte content of a unit of fresh whole blood is approximately 10^9 WBC/unit; this is reduced to 10^8 WBC/unit with the removal of the buffy coat layer. Leucocyte reduction using either filtration or apheresis will reduce the leucocyte content to 1 – 5x10^6 WBC/unit.

There is good evidence that transfusion of LD blood products results in reduction in febrile non-haemolytic transfusion reactions (FNHTR), reduces platelet refractoriness in patients receiving multiple platelet transfusions and reduces CMV transmission in susceptible patients (neonates, transplant). The evidence for reducing the risk of bacterial infection or cancer recurrence post resection is inconsistent, as is that for reducing short term mortality post transfusion except possibly in cardiac surgery and critically ill patients. There is no evidence for a reduction in the risk of reactivation of viral infections such as HIV and no change in survival for these patients.

LD should ideally happen pre-storage as leucocytes may fragment during storage and these fragments pass though the filters used for LD, in addition inflammatory cytokines which contribute to the FNHTR accumulate in stored blood and these will not be reduced by post-storage depletion. If the SANBS/WPBTS were to universally LD all products pre-storage their costs would increase by approximately 25% making the organisations unviable and leading to probable collapse of the South African service. The policy of the SANBTS (as per Clinical Guidelines 2014) is:

- All standard red cell concentrates are buffy coat depleted
- Random donor platelet concentrates are prepared from buffy coats
- Single donor platelet concentrates collected by apheresis must incorporate a leucocyte depletion process (standard practice with current apheresis technology)
- It is recommended that the following patients receive leucocyte depleted components:
- Patients on chronic transfusion regimens
  - Those at risk for CMV infection (all transplant patients)
  - Infants <1 year old
  - Critically ill, cardiac surgery and trauma patients (particularly those requiring massive transfusion
- Pre-storage (<48 hours after donation) leucocyte depletion in blood processing laboratories is recommended. If this is unobtainable the freshest components available may be filtered in the blood bank for immediate use (24 hour expiry). Bedside leucocyte depletion filters are not recommended unless neither of the former 2 options is available.

IRRADIATED BLOOD

Transfusion associated graft versus host disease (TA-GvHD) is an extremely rare but often fatal complication which may follow the transfusion of lymphocyte containing blood components. As with Graft vs Host Disease seen with stem cell transplant, cellular damage to the host skin, thymus, gastro-intestinal tract, liver and spleen result, in addition bone marrow hypoplasia is a feature that is specific to TA-GvHD. TA-GvHD has been reported following transfusion of whole blood, packed red cells, platelets and granulocytes but not FFP, cryoprecipitate or fractionated products.

While there is preliminary evidence that leucocyte depletion may reduce TA-GvHD the only means to prevent this complication is by gamma-irradiation of the product before transfusion. Red cell concentrates can be irradiated up to 14 days after collection and stored for a further 14 days without significant loss of viability.

Gamma irradiation of red cells leads to an accelerated leakage of potassium and an increase in extracellular levels of potassium. Hyperkalaemia may be a potential complication in rapid large volume transfusions such as intrauterine transfusion or neonatal exchange transfusion. Provided the unit is less than 5 days old this complication is unlikely. In the absence of fresh blood washing of cells will reduce the potassium content pre-transfusion.

Irradiated blood is recommended for:

- Blood components donated by blood relatives
- Intrauterine transfusion (IUT)
- Exchange transfusion (ET) following IUT
- Recommended for all ETs, provided this does not unduly delay the ET
- Platelets transfused in utero for alloimmune thrombocytopenia. Red cells and platelets transfused up to 6 months after the expected date of delivery should also be irradiated
- Lymphocyte immunodeficiency syndromes (incl. thymic hypoplasia)
- All recipients of allogeneic haemopoietic stem cell transplantation (HSCT) – from time of initiation of conditioning regimen. This should continue while the patient is on GvHD prophylaxis or lymphocytes are >1 x 109/l
- Patients undergoing autologous stem cell harvesting – until there is evidence of haematopoietic engraftment and lymphoid reconstitution
- Hodgkin lymphoma
- Treatment with purine analogues (fludarabine, cladribine, deoxycoformycin)
- Anti-thymocyte globulin (ATG) for severe aplastic anaemia

COMPATIBILITY TESTING

When group & screen is ordered testing for ABO and Rhesus groups and a screening test for clinically significant antibodies (Kell, Duffy, Kidd etc) is done. Blood is only cross matched at this stage if clinically significant antibodies are detected.

In infants <4 months of age antibodies derive from maternal blood and so a sample of maternal blood can be used for cross-matching. The development of antibodies to red cell antigens is very uncommon in the first 4 months of life and certain blood transfusion services recognize this and are happy to waive the requirement of repeated blood sampling for ongoing transfusions for infants in this age group.

Blood group	Antigens	Antibodies	Can give blood to	Can receive blood from
AB	A and B	None	AB	AB, A, B, O
A	A	B	A and AB	A and O
B	B	A	B and AB	B and O
O	None	A and B	AB, A, B, O	O

Table 7: Major donor recipient blood grouping

Cross-matching of blood determines whether the recipient has pre-formed antibodies against any antigens on the donor's cells and whether a donor's blood is compatible with that of an intended recipient. A complete cross-matching process takes about an hour. If blood is needed in an emergency ABO and Rh type-specific blood can be requested. The blood bank will continue the cross matching process in this case while the transfusion is in progress but the clinician accepts the risk of administering incompletely cross-matched product.

Universal donor blood (O- or O+ for males) may be kept in appropriate conditions in institutions where blood may be required in less time that it takes to procure from a blood bank.

ADMINISTRATION FACTORS

Red blood cells, whole blood, cryoprecipitate, FFP and WPBTS VIAHF (Factor VIII concentrate) are administered through a standard blood recipient set (170 - 240 µm mesh filters) to prevent the transfusion of clots or coagulation debris. The use of micro aggregate (40 µm) filters is not recommended.

A platelet giving set should be used for platelet transfusion. Use of a standard blood administration set will result in greater loss of the available platelets due to a larger surface area for adhesion.

Blood transfusions must be completed within 6 hours of entry of the pack. Platelets should be given over 20 – 30 minutes.

Rapid transfusion of cold product may result in hyperthermia however care should be taken with rewarming as overheating of the blood can cause extensive haemolysis.

The Physiology of Ageing

Dr Adalbert Ernst

UCT Dept of Anaesthesia & Perioperative Medicine

Introduction

The development of sexual reproduction in eukaryotic cells over a billion years ago had a payoff: once two organisms had combined their genetic material to produce offspring, they were, evolutionary speaking, disposable, and there was no need to preserve individuals indefinitely once they had passed on their DNA. As such, the progressive decline in physiological function that eventually results in death is a normal phenomenon present in complex multicellular lifeforms. *Homo sapiens*, despite being one of the longest-lived of all mammals, is no exception.

These notes will concentrate on the key physiological changes that occur as humans age.

What Is Ageing?

Ageing represents *an accumulation of changes in an organism over time, manifesting as an incremental fall in physiological performance that eventually results in total loss of function (death)*. In humans, this encompasses not only physical but also psychological and social factors. It is thought that one of the reasons humans are relatively long-lived is that we have a complex social structure that is not merely dependent on passing on DNA. Indeed, longitudinal studies have shown that individuals with active social bonds and meaningful relationships live longer than those who do not.

Senescence refers explicitly to the **biological** processes involved in ageing.

Cellular senescence describes the process whereby cells lose the ability to undergo mitosis. (Normal human cells lose the ability to divide after roughly 50 generations, the so-called Hayflick Limit).

Frailty refers to the increased vulnerability to adverse outcomes as a consequence of ageing.

Theories of Ageing

There is no single unified theory of ageing. Even within the same species, different organisms will age at different rates. The current view is that ageing is a **heterogeneous** process of **gradual physiological decline**, resulting from an interplay of various mechanisms at molecular, cellular, organ, and systemic level. This is further influenced by environmental and epigenetic factors such as nutrition, infection, and trauma.

Broadly speaking, there are two major schools of thought regarding ageing.

a) The **programmed theory** of ageing supposes that, over time, there is a specific and predictable decline in homeostasis and natural defence mechanisms encoded into the genome and influenced by epigenetic factors.

b) The **error theory** proposes that the body undergoes progressive damage due to various environmental insults, particularly due the production of reactive oxygen species (ROS).

The key elements involved can be summarized as:

1. Progressive reliance on homeostatic reserves (“homeostenosis”)
2. Reduced redundancy of function (reduced reserve)
3. Impairment / reduction in repair mechanisms

Proposed molecular mechanisms involved in ageing

A clear understanding of the molecular mechanisms involved remains elusive and is thought to be an interplay of several factors, including:

Telomere shortening
Telomeres are repeated nucleotide sequences that “cap” each end of a chromosome, protecting it from damage. In most human cells, telomere length reduces with each division, because the enzyme responsible for replacing the lost ends of telomeres (“telomerase”) falls to suboptimal levels, eventually becoming unexpressed. The telomere hypothesis of ageing points out that many age-associated changes are caused by cellular senescence induced by crucially shortened telomeres.

Oxidative stress
Reactive oxygen species (ROS) such as singlet oxygen and hydrogen peroxide, generated during mitochondrial energy production, cause damage to DNA, proteins, and lipids. Superoxide dismutase (SOD) is a free-radical scavenger. There is evidence that SOD activity is increased in some long-lived organisms. Caloric restriction is known to result in longer life, reflecting reduced ROS production from decreased metabolism. Furthermore, the amino acid cysteine is essential for regenerating glutathione, one of the so-called “master antioxidants”. There is some evidence that a cysteine rich diet or even n-acetylcysteine supplementation may attenuate oxidative stress by ensuring optimal glutathione levels.

DNA damage
DNA may be damaged due to oxidative stress (see above). This includes damage to both nuclear and mitochondrial DNA (mt-DNA). The very act of DNA replication may also introduce errors into the cellular milieu with missense or nonsense mutations resulting in chain termination, or, occasionally, productions of abnormal peptides or proteins which may contribute to structural macromolecular changes.

Tumour suppression as a cause of ageing
There is growing evidence that cancer and ageing may represent two sides of the same coin. Malignancy is held in check by inbuilt, cellular tumour-suppressing mechanisms that limit unregulated cell division. However, these anti-cancer pathways also drive cellular senescence and apoptosis, which may influence age-related decline. The p53 pathway is a particularly crucial suppressor of malignant change in multicellular organisms and it is thought it actively promotes cellular senescence to prevent cancer. Animal and in vitro models have shown that increased p53 activity leads to phenotypes consistent with accelerated ageing.

Stem cell depletion
An attractive hypothesis is that stem-cell depletion in mitotically active tissues leads to ageing. This theory posits that many of the processes described above (DNA damage, telomere shortening, and mitochondrial dysfunction) result in widespread stem-cell depletion, resulting in decreased regenerative ability.

Reactivation of Developmental Programmed Senescence (DPS)
DPS is a recently discovered pathway that is distinct from the senescence induced by tumour suppression, and is active in developing embryonic tissue. It is normally self-limiting, does not involve tumour suppressor genes, and is considered essential to normal embryonic development. However, there is in vitro evidence that DPS may be induced or “switched on again” by increased levels of cytokines such as TGF (transforming growth factor) in an inflammatory milieu.

Epigenetic factors
One should not forget about the external environment! The effects of lifestyle (diet, exercise, habits, and exposure to toxins) have a significant impact on physiology. Exercise increases free-radical scavenging ability, while smoking is known to accelerate ageing by depleting telomere length.

Gender
It is well known that women generally live longer than men, even when accounting for behaviour and habits. This may be to a variety of reasons: (1) As males only have one X chromosome, they lack a backup chromosome that can be referenced during DNA repair during cell replication. (2) Testosterone is a relatively pro-oxidant hormone compared to oestrogen, and induces inflammatory changes in tissue, particularly during puberty. (3) There is even speculation that the heart-rate increase that occurs during the menstrual cycle induces metabolic adaptation equivalent to moderate exercise.

Other theories
Growth Hormone: There is some evidence that reduced growth hormone (GH) signalling correlates with increased lifespan (in mice) despite claims that GH supplementation may prolong life.
IGF-1 signalling: Long-lived mice and mice subjected to dietary restriction have reduced insulin and insulin-like growth factor (IGF-1) levels, and there is evidence for a genetic mechanism whereby increased insulin sensitivity confers longevity (independent of the negative effects of increased insulin resistance e.g. diabetes)

KEY PHYSIOLOGICAL CHANGES INVOLVED IN THE AGEING PROCESS

Cardiovascular System

The chief changes involve impairment of mechanics and contractile efficiency:

1. Arterial wall thickening, increased smooth muscle tone, and changes in the vascular matrix (loss of elastin and collagen) lead to “stiffening” or rigidity the vessels (especially large, elastic arteries)
2. This, in turn, results in elevated systolic arterial pressures, increased systemic vascular resistance (SVR), and increased afterload.
3. Cardiac work and oxygen demand increase.
4. Isolated systolic hypertension is common. Left ventricular hypertrophy (LVH) with narrowing of the left ventricular outflow tract (LVOT) may result as the left ventricle contracts more forcefully into the stiffened aorta.
5. Hypertrophy of cardiac myocytes occurs secondary to increased afterload, lengthening contraction time.
6. Occult diastolic dysfunction is common due to delayed ventricular relaxation.
7. Early diastolic filling rate decreases with age, but this buffered by an increase in late diastolic filling. This partly explains the correlation between increased age and left atrial size, itself increasing the likelihood of atrial fibrillation (AF).
8. A progressive decline in atrial pacemaker cells (>50%) leads to a decrease in intrinsic automaticity and increased risk of atrial arrhythmias.
9. Decreased stroke volume results in decreased cardiac output.
10. Both aortic arch and carotid sinus baroreceptor function is impaired (delayed) with attenuation of the heart rate response to changes in arterial pressure. This, combined with age-related autonomic dysfunction (see under Nervous System) compromises haemodynamic homeostasis. Blood pressure becomes more labile, with increased incidences of postural hypotension, post-prandial hypotension, sinus node depression, carotid sinus syndrome, and syncope. It is important to appreciate that the baroreceptor deterioration is multifactorial (reduced arteriolar compliance, blunted transduction of stretch signals, altered central neural processing, altered baseline efferent autonomic outflows, and damped end-organ responsiveness).
11. The heart rate response to exercise is attenuated. Despite this, regular aerobic exercise improves physiological functional capacity, VO_2 max, aerobic capacity, arterial compliance, and endothelium-dependent dilatation. Exercise also is helpful in raising free oxygen radical scavenging capacity, regenerating endothelium and intima-media wall thickness.
12. There is reduced compliance of the venous system, with decreased ability to buffer changes in volume.

Respiratory System

1. Loss of elasticity in the bony thorax and airways leads to easily collapsible alveoli, leading to a decreased surface area available for alveolar gas exchange.
2. The concomitant loss of muscle mass typical with ageing complicates the situation; muscles involved in respiration are weakened. There is increased chest wall rigidity. As a result, FRC is increased while total lung capacity remains unchanged.
3. Closing capacity (CC) encroaches on tidal volume resulting in V/Q mismatching and reduced PaO_2. CC reaches FRC by the mid-forties in the supine position, and by 66 when upright. Closure of lung bases redistributes inspired gas to apical areas of the lung which are underperfused (increasing dead space) while dependent areas of the lung are underventilated (increased shunt). Shunt tends to be the predominating factor.

4. The age related-effect on normal arterial oxygen tension may be estimated using the following formula:

    i. PaO_2 (kPa) = 13.3 – (age/30)

    ii. Aside from V/Q mismatching, the reduced PaO_2 reflects altered lung mechanics, diminished diffusion capacity, and reduced cardiac output resulting in increased oxygen extraction and reduced mixed venous oxygen tension.

5. Compliance across the lung is reduced in a non-uniform fashion: some regions still empty normally, while passive exhalation is slowed in others.
6. Lung expansion is less effective as respiratory rate increases, further exacerbating V/Q mismatching.
7. Centrally, there is a blunted CNS response to both hypoxia and hypercapnia

Nervous System

1. Neural density decreases: 30% of brain mass is lost by age 80, primarily grey matter.
2. A reduction in neurotransmitters (including catecholamines, serotonin, and acetylcholine) leads to subsequent deleterious effects on mood, memory and motor function.
3. Cortical binding sites for serotonin, GABA and catecholamines become depleted.
4. Signal transduction in the brainstem and spinal cord progressively declines.
5. There is loss of motor, sensory and autonomic fibres in the PNS, with reduction in both afferent and efferent conduction velocity
6. Axons innervate fewer muscle cells, leading to denervation and muscle atrophy.
7. The dominant autonomic tone gradually becomes sympathetic as baseline parasympathetic outflow decreases. Despite this, the response to beta-adrenergic stimulation is blunted.
8. Circadian rhythm alters as we age, typically leading to fewer hours of effective sleep, attenuated natural body temperature rhythms, and early awakening. Individuals may compensate by going to bed earlier. Cognition is typically best in the morning and worsens during the day.

Endocrine System

1. Alterations in signal transduction may decrease the ability of target organs to respond to target hormones.
2. Concentrations of many hormones change with age, but with little demonstrable clinical relevance.
3. A higher serum ADH concentration may result from altered baroreceptor function, placing the elderly at risk for hyponatraemia.
4. Carbohydrate intolerance increases through a variety of mechanisms (mainly increased adiposity, loss of muscle mass and decreased fitness).
5. Reduced Vitamin D synthesis in skin leads to decreased calcium absorption and predisposes to osteopenia and osteoporosis.
6. Menopause in females predisposes to osteoporosis. Testosterone secretion in males falls progressively after the mid-forties and may be significant, leading to the so-called partial androgen deficiency of the ageing male (PADAM).
7. Insulin resistance may reflect ageing and apoptosis of pancreatic islet beta-cells, predisposing to Type 2 Diabetes.

Renal Changes

1. Loss of renal mass begins from the fourth decade, and is predominantly cortical with relative sparing of the medulla.
2. The number of glomeruli is reduced. Sclerosis and diminished lobulation of existing glomeruli lead to reduced filtration areas and hence an age-related decline in GFR.
3. Tubular atrophy and fibrosis is common.

4. Even in the absence of diabetes, hypertension or chronic renal disease, glomerular basement membrane permeability increases, leading to some degree of microalbuminuria and proteinuria.
5. Renal blood flow declines progressively by 10% per decade after the age of 30, mainly in the cortex with a relative increase in flow to the juxtamedullary region.
6. Impaired vasodilatation of the afferent arteriole due to an imbalance of humoral factors.
7. Creatinine clearance is influenced by nutritional status, protein intake, muscle mass, body weight, gender, and ethnicity. As muscle mass generally decreases with ageing and urinary creatinine excretion decreases, the result is that the declining GFR is accompanied by lower rises is serum creatinine (relative to younger patients).

Immunological Changes

1. Senescence of immune cells predisposes to infection and delayed recovery.
2. Both innate and acquired immunity are affected.
3. Macrophage, B- and T-cell function decline, resulting in impaired humoral and cellular responses.
4. Dendritic cell number decreases, but function appears unimpaired.
5. Complement system activation is blunted in response to inflammation.
6. A reduced capacity to generate inflammatory mediators such as TNF-alpha, interleukin-1, and nitric oxide predispose to reactivation of dormant infections (e.g. shingles, TB) and susceptibility to new infections.
7. Increased autoimmunity occurs, with an increase in autoantibodies (both specific and non-specific.)

Gastrointestinal System

1. With age, there is desynchronization between the contraction and relaxation mechanisms in the oropharynx and oesophagus during swallowing.
2. Impaired taste and smell, faster antral filling and early satiation (the latter due to elevated cholecystokinin and reduced ghrelin levels) may predispose to age-related anorexia.
3. HCl and pepsin secretion are decreased, causing a small rise in gastric pH.
4. Calcium absorption is impaired, but it is not known whether this is due to vitamin D deficiency (common in the elderly) or a primary malabsorption process.
5. Prolonged gut transit time predisposes the elderly to constipation.

The Skin

1. Impaired barrier function, reduced epidermal cell turnover, and a reduction in keratinocyte and fibrinolytic result in increased fragility and loss of elasticity; this occurs in tandem with sunlight-induced photo-ageing which becomes more prominent over the years.
2. Reduced vascularity results in fibrosis and skin atrophy.
3. Parallel immune senescence renders skin more vulnerable to infection and neoplasia.

Musculoskeletal System

1. Sarcopenia (loss of muscle strength) reflects a 30% decline in muscle mass from the 3rd to the 8th decade with reduced cross-sectional area.
2. Mainly Type II (fast-twitch) muscle fibres are reduced, significantly reducing VO_2 max and force of contraction. Habitual aerobic exercise counteracts this. Slow-twitch fibres are relatively spared.
3. Loss of elasticity in joints (due to changes in collagen structure) is near universal.
4. Males over 50 lose bone at a rate of 1% per year, whereas loss is 2-3% in women after menopause. The loss in bone mineral density predisposes to osteopenia and osteoporosis with an increased risk of fractures.

Energy and Thermoregulatory Changes

1. Reduced total energy expenditure and physical activity may reduce basal metabolic rate (BMR) by as much as 20%.

2. Gluconeogenesis, fatty acid oxidation and Na^+/K^+/ATPase activity are all reduced with altered mitochondrial membrane permeability.
3. The threshold for detecting changes in skin temperature rises with ageing. Decreased vasomotor responses result in skin being less able to conserve or lose heat appropriately. Shivering threshold and shivering effectiveness are impaired. The net effect is that the elderly are at higher risk from adverse effects in both hot and cold environments (heat exhaustion/dehydration and hypothermia, respectively.)

References

1. Kaeberlein, M., 2007. Molecular basis of ageing: Meeting Report, *European Molecular Biology Organization Reports*, Vol 8(10)
2. Liu, J., 2014. Molecular mechanisms of ageing and related diseases. *Clin Exp Pharmacol Physiol* 41, 445-458
3. Maguire, S and Slater, B., 2013. Physiology of Ageing. *Anaesthesia and Intensive Care Medicine* 14(7)
4. Navaratnarajah A and Jackson, S., 2013. The physiology of ageing. *Medicine* 41(1)

Chronic Pain Physiology

Dr Tory Madden

Postdoctoral research fellow- Pain group
UCT Depts of Anaesthesia & Perioperative Medicine, and Psychiatry & Mental Health

Modern pain scientists understand pain to be a perceptual experience that is generated in response to perceived threat (for review, see Moseley and Butler 2017). In other words, the brain generates this experience that is pain when its 'best guess' is that bodily tissue is at risk of damage. This inference (that bodily tissue is under threat) relies on a plethora of information from both internal and external sources. These sources could include interoceptive information, knowledge of the context (based on prior first-hand experience or vicarious learning), visual information, somatosensory signalling, body position, information provided by a healthcare clinician, and many other sources.

In the optimal system, pain occurs *before* tissue damage and motivates a change in behaviour that *prevents* tissue damage. A typical example of this is the pain one might feel when climbing into a bath that is too hot – usually one feels pain and gets out again without sustaining tissue damage. That is helpful or *adaptive* pain, because it prevents tissue damage, thus promoting survival.

A functional definition of chronic pain is *pain that has outlived its usefulness*. Formal definitions used for research usually run along the lines of pain on most days (or every day) for three months or longer. The principle is that there is no longer tissue damage, but pain persists. In some cases, there is tissue damage but it is not sufficient to explain the pain. The pain is unhelpful because it does not protect tissue that needs protection. In many people, this chronic pain interferes with their engagement in life, which makes it *maladaptive*.

The brain as an integrator of information

It is helpful to think of the brain as responsible for integrating all the information available to it and generating our perceptual experiences to optimise our survival and engagement with life (Tabor, Thacker et al. 2017). The brain only has access to a certain amount of information, and much of this processing is not done at a conscious level. The prerogative to protect seems strong enough that we get 'stuck' with a bias towards perceiving threat more commonly than we get stuck with a bias towards perceiving safety (think of, e.g., generalised anxiety disorder, chronic pain, social phobias). This is what seems to happen in chronic pain – the brain gets 'stuck' perceiving threat, and so generates pain in response to information that otherwise wouldn't be sufficient to elicit pain.

In clinical reasoning, it can be helpful to think of all the pieces of information that the brain is receiving that could skew its conclusion towards an inference that bodily tissue is under threat or towards the inference that bodily tissue is safe.

Information sources

Historically, research and clinical treatment of pain have focused on trying to modify nociceptive signalling. But nociceptive signalling is only one source of information – and, critically, nociception is neither sufficient nor necessary for pain. There is also a wealth of other information that we have an opportunity to influence. So, while it is important to be aware of the changes along the neuraxis that alter afferent nociceptive signalling, it is equally important to be aware of the influences of other information sources.

Importantly, the terms 'pain receptors' and 'pain pathways' are inaccurate and interfere with the effort to understand and treat pain. Although some are struggling to change their use of this inaccurate language, it is important to use accurate terms and, when reading older papers, to rephrase the terms to one's self – by 'pain receptors' people usually mean nociceptors, and by 'path pathways' people usually mean nociceptive pathways. Similarly, using inaccurate language with patients is also unhelpful and interferes with the priority of helping people to understand and manage their own pain. 'Danger' or 'warning' receptors/messages are more helpful terms and more closely represent the current understanding of nociception and pain.

Changes along the neuraxis associated with persistent pain

Peripheral sensitisation

- Note that a neural fibre may be considered 'nociceptive' if it has a high response threshold – i.e. it only responds to fairly intense stimulation. Therefore, when we speak about nociception, we are speaking of signalling about *intense* stimulation that could be mechanical, thermal or chemical in nature.
- When tissue is damaged, various components of the inflammatory soup can influence afferent signalling:
    - From damaged cells: potassium, histamine, serotonin, bradykinin, ATP.
    - Synthesised at the site of tissue damage: prostaglandin, leukotriene
    - Released by nociceptors: Substance P
- Processes:
    - Reduced firing threshold of nociceptors
    - Increased responsiveness of nociceptors
    - Activation of 'silent' nociceptors
- Note the role of antidromic activity in C-fibres, driving neurogenic inflammation
- Responsible for primary hyperalgesia (increased pain to a stimulus that is normally painful; restricted to the site of tissue damage)
- Net result is that afferent signalling along nociceptive pathways is increased.
- Peripheral sensitisation is *normal and adaptive* when it is an acute-phase response to tissue damage.

Central sensitisation in the dorsal horn

- Nociceptive and non-nociceptive information arrives at the dorsal horn and most of the fibres carrying this information meet synapses at the same level or nearby levels.
- Synaptic transmission is influenced by:
    - Top-down control: facilitation and inhibition of synaptic transmission is regulated by descending modulatory influences from the brain (RVM and PAG are prime players here). In persistent pain, the 'balance' of downward modulation of dorsal horn synapses shifts, resulting in increased facilitation and/or decreased inhibition.
    - The amount of afferent signalling arriving at the synapses. Repeated stimulation has the effect of temporal summation (increased activity in second order neurons); stimulation across a wide area results in spatial summation. Repeated activation of C-fibres leads to the central release of excitatory amino acids, glutamate (excitatory; NMDA/AMPA) and peptides such as CGRP and substance P (NK-1). In addition, prolonged activation of NMDA receptors results in changes in gene expression that lead to reduces firing thresholds of second order neurons.
- Heterotopic facilitation is a kind of cross-modal facilitation effect via which increased activity in some classes of primary afferent neurons results in sensitisation of other classes of neurons and/or sensitisation of neurons that carry information from other (nearby) areas. This manifests as allodynia (pain to a stimulus that is normally non-painful) and/or secondary hyperalgesia (increased pain to a stimulus that is normally painful; in a region adjacent to the site of tissue damage – i.e. in *undamaged* tissue).
- The receptive fields of second order neurons are plastic and their precision is maintained by inhibitory signalling from higher centres. When this inhibitory signalling changes (see above – top-down control), the receptive fields of the second order neurons are no longer maintained as well, and they may widen so that these neurons receive input from primary afferent neurons they didn't previously synapse with. Many second order neurons don't have the capacity to respond to non-nociceptive neurons, so the change in receptive field will only affect their activity if they now synapse with more nociceptive neurons. In contrast, 'wide dynamic-range' secondary neurons are quite versatile and have the capacity to respond to nociceptive or non-nociceptive input. If these neurons' receptive fields shift they may start to respond to innocuous input from non-nociceptive sources. The result is that the brain receives input via a WDR pathway that is usually reserved for nociception, and so it may interpret that signalling as threatening – even if the signal was initially a response to a light touch or another innocuous event.

- Microglia (normally functioning as the macrophages of the CNS, detecting and removing pathogens, etc.) appear to play an important role in limiting inhibition at the dorsal horn – specifically, microglia seem to be activated by peripheral nerve injury and unmask previously inhibited synapses in the dorsal horn.
- Astrocytes form part of the tripartite synapse and influence synaptic function continually in the normal state (respond to neurotransmitters by releasing glial substances into the synapse). Astrocytes respond to synaptic transmission by releasing glutamate (increasing NMDA activation), prostaglandin and pro-inflammatory cytokines, consequently increasing the signalling strength of synapses. They also increase neuronal AMPA expression. They therefore play an important role in upregulation of synaptic signalling in the CNS. Interestingly, astrocytes signal between themselves via 'calcium waves'.

Features in the brain (see Moseley and Flor 2012)

- It is reasonable to consider that the processes underlying changes in neuronal signalling at the dorsal horn probably affect signalling in the brain similarly. In general, there is a loss of inhibition and more widespread activation.
- The loss of cortical inhibition affects brain-held body maps in the homunculus – sometimes called homuncular 'smudging' – resulting in poor localisation or mis-localisation of somatosensory input.
- Studies that image brain activation during experimental stimulation, comparing people with chronic pain and people without pain, have found a generally increased level of activation of areas linked to emotion in those with chronic pain when somatosensory stimulation is provided (e.g. Hashmi, Baliki et al. 2013).
- Beware of the idea of a 'pain matrix': it implies the existence of brain activation patterns that are specific to pain – a specificity that has not been verified. It seems more likely that the patterns of activation in the brain seen on painful stimulation actually reflect salience, or the relevance/importance of the stimulation: the activation diminishes with repetition of the painful stimulation, and similar activations are seen with non-painful but similarly salient stimulation.

Changes in other systems that may affect nociception and/or pain

Immune system:
- Upregulation of glial activity – microglia and astrocytes influence synapses (tripartite synapse)
- The balance in cytokines can influence nociception; a shift in cytokine balance towards a more pro-inflammatory profile is likely to upregulate nociceptive processes.

Endocrine system, stress-related changes:
- Cortisol (regulates healing) may be diminished ('burned out')
- Corticotropin releasing hormone levels may be increased and can promote nociception via several different routes – most notably by upregulating locus caeruleus activity, which results in increased arousal, and by increasing pro-inflammatory cytokine production — or decrease nociception by promoting opioid release.

Autonomic nervous system:
- The vagus nerve has attracted attention as a source of anti-nociceptive activity. There are complex interactions between the autonomic nervous system and the somatosensory system, and the effects of the extremes of arousal (either very high arousal e.g. escape, or very low arousal, e.g. somnolence) seem to be generally to reduce pain. This can be understood at an anti-nociceptive level or at a behavioural prioritisation level (i.e. in high arousal, the priority is to escape a BIG threat rather than to protect tissue from a smaller threat).

Emotions and cognitions:
- These are an important influence on a person's overall sense of threat or safety (Wiech and Tracey 2009). There may be a disconnect between consciously stated beliefs or thoughts and ideas that may not be subject to conscious consideration, and any of these have the potential to shift the inferential balance towards or away from pain. Even in the laboratory, manipulation of the implicit threat value of an experimental somatosensory stimulation can determine whether or not that stimulation elicits pain (e.g. Wiech, Lin et al. 2010).

References

Hashmi, J. A., M. N. Baliki, L. Huang, A. T. Baria, S. Torbey, K. M. Hermann, T. J. Schnitzer and A. V. Apkarian (2013). "Shape shifting pain: chronification of back pain shifts brain representation from nociceptive to emotional circuits." Brain **136**(9): 2751-2768.
Moseley, G. and D. Butler (2017). "Explain pain supercharged." Adelaide, Australia: Noigroup publications.
Moseley, G. L. and H. Flor (2012). "Targeting Cortical Representations in the Treatment of Chronic Pain: A Review." Neurorehabilitation and Neural Repair **26**(6): 646-652.
Tabor, A., M. A. Thacker, G. L. Moseley and K. P. Körding (2017). "Pain: A Statistical Account." PLOS Computational Biology **13**(1): e1005142.
Wiech, K., C. Lin, K. H. Brodersen, U. Bingel, M. Ploner and I. Tracey (2010). "Anterior Insula Integrates Information about Salience into Perceptual Decisions about Pain." The Journal of Neuroscience **30**(48): 16324-16331.
Wiech, K. and I. Tracey (2009). "The influence of negative emotions on pain: Behavioral effects and neural mechanisms." NeuroImage **47**(3): 987-994.

The Pharmacology of Amiodarone and Digoxin as Antiarrhythmic Agents

Dr Adri Vorster

UCT Department of Anaesthesia & Perioperative Medicine

The heart contains pacemaker, conduction and contractile tissue. Cardiac arrhythmias are caused by either enhancement or depression of cardiac action potential generation by pacemaker cells, or by abnormal conduction of the action potential. The pharmacological treatment of arrhythmias aims to achieve restoration of a normal rhythm and rate.

The resting membrane potential of myocytes is around -90 mV, with the inside of the membrane more negative than the outside. The main extracellular ions are Na^+ and Cl^-, with K^+ the main intracellular ion. The cardiac action potential involves a change in voltage across the cell membrane of myocytes, caused by movement of charged ions across the membrane. This voltage change is triggered by pacemaker cells. The action potential is divided into 5 phases (figure 1).

Phase 0: Rapid depolarisation

Duration < 2ms

Threshold potential must be reached (-70 mV) for propagation to occur

Rapid positive charge achieved as a result of increased Na^+ conductance through voltage-gated Na^+ channels in the cell membrane

Phase 1: Partial repolarisation

Closure of Na^+ channels

K^+ channels open and close, resulting in brief outflow of K^+ and a more negative membrane potential

Phase 2: Plateau

Duration up to 150 ms

Absolute refractory period – prevents further depolarisation and myocardial tetany

Result of Ca^{++} influx through voltage-sensitive L-type Ca^{++} channels, K^+ efflux and Cl^- influx, with a near balance of ion movement

Phase 3: Repolarisation

Membrane potential returns to resting value

Relative refractory period – supra-normal stimulus required for contraction

Result of increased K^+ conductance and closure of Ca^{++} channels, with a net outward positive current

Phase 4: Resting potential

The resting membrane potential is mainly determined by K+ equilibrium and the Na^+/K^+ ATPase pump (figure 2) maintains ionic concentration gradient and membrane potential at approximately -90 mV by exchanging 3 Na^+ for 2 K^+ using ATP

The sodium-calcium exchanger also trades 1 Ca^{++} from the cell for 3 Na^+ into the cell, using ATP

Figure 1. Myocyte action potential

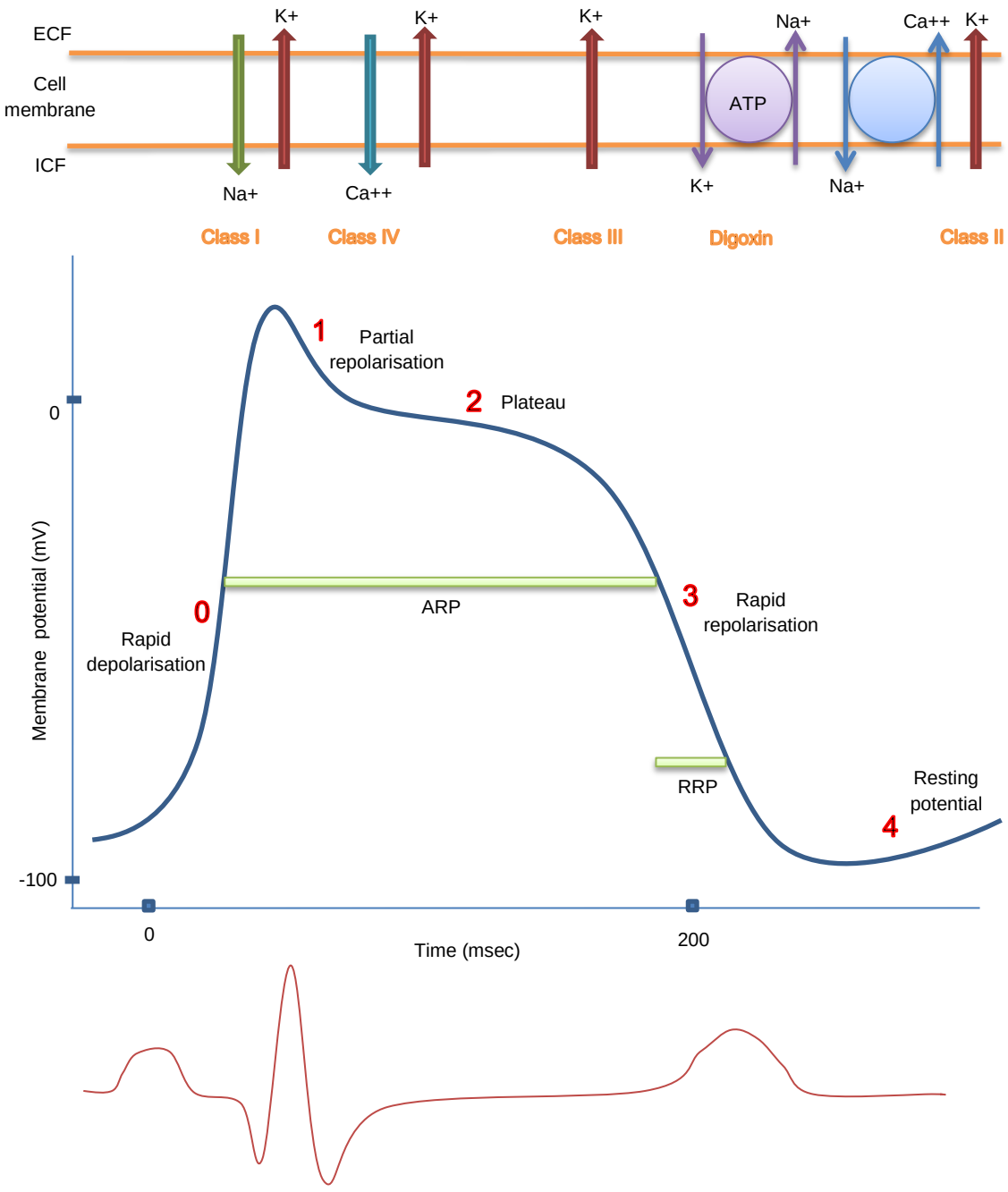
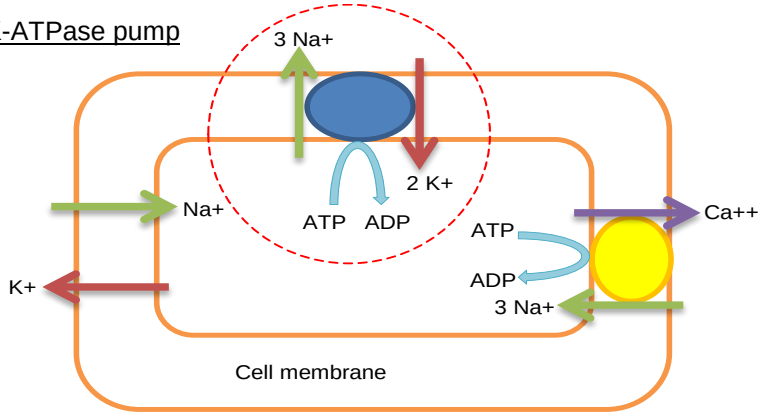


Figure 2. Na-K-ATPase pump



Traditionally antiarrhythmic agents have been classified according to the Vaughan-Williams classification (table 1). This classification does not include agents like digoxin and adenosine.

Some agents also exhibit characteristics that place them in more than one category, e.g. amiodarone (classes Ia, II, III, IV) and sotalol (classes I, II, III). Other methods of classification may include type of rhythm treated or site of action (table 2).

Amiodarone:

Amiodarone is an agent used in the treatment of supraventricular and ventricular tachyarrhythmias, as well as Wolff-Parkinson-White Syndrome. It is classified as a Vaughan-Williams class III antiarrhythmic agent, but it is a complex drug with multiple actions. It blocks K^+ channels, therefore slowing repolarisation and increasing the duration of the action potential.
It also shows sodium and calcium channel blocker properties, and is a non-competitive ß-adrenergic inhibitor. Amiodarone chemically resembles thyroxine and some of its pharmacological effects may be explained by its binding to thyroid receptors.
Amiodarone has multiple side effects and it will affect most patients on chronic treatment. Most side effects are reversible with cessation of treatment, but some side effects have a significant mortality rate e.g. pneumonitis, hepatitis, exacerbation of asthma and congestive cardiac failure.

Digoxin:

Digoxin is a cardiac glycoside that is extracted from the leaves of the foxglove (*Digitalis lanata*) and is widely used in the treatment of atrial fibrillation and atrial flutter, and also in the treatment of cardiac failure.
Digoxin has positive inotropic and negative chronotropic activity due to direct and indirect actions on the heart.
Directly it inhibits the Na^+/K^+ ATPase pump, leading to increased intracellular Na^+ and decreased intracellular K^+ concentrations. The raised intracellular Na^+ concentration leads to reversal of the action of the sodium-calcium exchanger, increasing the exchange of intracellular Na^+ for extracellular Ca^{++}. This results in increased availability of intracellular Ca^{++}, with a positive inotropic effect. The refractory period of the AV node and bundle of His is increased and conduction is therefore slowed down.
Indirectly it increases the release of acetylcholine at myocardial muscarinic receptors. This also increases the refractory period of the AV node and bundle of His, slowing conduction. In addition it may improve baroreceptor sensitivity in patients with cardiac failure and increases renal blood flow.
Digoxin has a narrow therapeutic index and side effects are not uncommon. Cautious dosage determination and monitoring of therapeutic levels are essential. Side effects are also more common in patients with hypokalaemia, as digoxin competes with K^+ ions for the same binding site on the Na^+/K^+ ATPase pump.

The basic pharmacological characteristics of amiodarone and digoxin are summarised in table 3.

Classification of antiarrhythmic agents

Table 1.

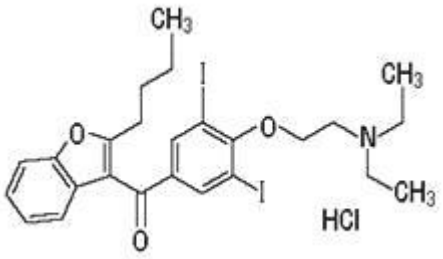
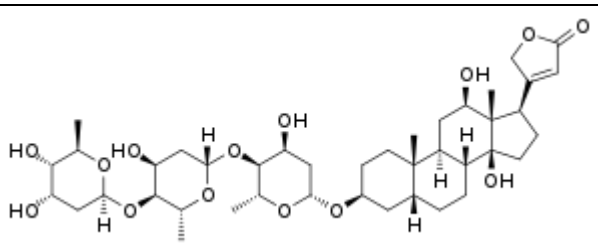
Vaughan-Williams	Site	Drug examples	Action	AP	HR	PR	QRS	QT	EA	CV	RP
Class Ia	Na channel blockade	Quinidine Procainamide Disopyramide	Slow depolarisation ++	↑		↑	↑	↑	↓	AV↑↓ CS↓	↑
Class Ib	Na channel blockade	Lidocaine Phenytoin Mexiletine Tocainide	Slow depolarisation +	↓	↔	↔	↔	↔↓	↓	AV↔ CS↔↓	↑↓
Class Ic	Na channel blockade	Flecainaide Propafenone	Slow depolarisation +++	↔	↔	↑	↔↑	↔↓	↓	↓	↑
Class II	ß - adrenergic blockade	Propanolol Carvedilol Metoprolol	AV nodal blockade		↓	↑	↔	↔↓	↓	AV↓ CS↔	AV↑ CS↔
Class III	K channel blockade	Amiodarone	Slow repolarisation	↑	↓	↑	↑	↑	↓	↓	↑
		Sotalol Bretylium Dronedarone Ibutilide			↓	↑	↔	↑	↓	AV↓	↑
Class IV	Ca-channel blockade	Verapamil Diltiazem	AV nodal blockade		↓	↑	↔	↔	↓	AV↓ CS↔	AV↑ CS↔
Other		Digoxin			↓	↑	↔	↓	↑	AV↓ CS↔	AV↑ CS↓
		Adenosine				↑			↓	AV↓ CS↔	AV↑ CS↔
		Magnesium							↓		

AP=action potential, HR=heart rate, EA=ectopic automaticity, CV=conduction velocity, AV=AV node, CS=conduction system (His-Purkinje and ventricular), RP=refractory period

Table 2

Site of Action	Drug examples
Sinus node and Atrium	ß -blockers Digoxin Amiodarone Quinidine Procainamide Disopyramide Verapamil
Atrio-ventricular node	ß-blockers Digoxin Ca-channel blockers Flecainaide Propafenone
Accessory Pathway	Amiodarone Disopyramide Procainamide Flecainaide
Ventricle	Amiodarone Quinidine Procainamide Disopyramide ß -blockers Lidocaine Mexiletine Phenytoin

Table 3

	Amiodarone	**Digoxin**
Formula	C25H29I2NO3	C41H64O14
Chemical structure		
Chemical name	(2-{4-[(2-butyl-1-benzofuran-3-yl)carbonyl]-2,6-diiodophenoxy}ethyl)diethylamine	(3Beta,5B,12B)-3-[(0-2,6-dideoxy-B-D-ribo-hexopyranosyl-(1 to 4)-0- 2,6-dideoxy-B-D-ribo-hexopyranosyl-(1 to 4)-2,6 dideoxy-B D-ribo-hexopyranosyl)oxy]-12,14-dihydroxycard-20(22)-enolide
Description	Iodinated benzofuran derivative	Sterol lactone + sugar
Classification	Vaughan-Williams Class III antiarrhythmic agent (also exhibit actions of classes Ia, II, and IV)	Cardiac glycoside
Mechanism of action	**Blocks myocardial calcium, potassium and sodium channels** → prolongation of cardiac action potential and refractory period **Inhibits alpha- and beta-adrenergic receptors** → reduction in sympathetic stimulation, negative chronotropy, and ↓ myocardial oxygen demands **Vasodilation** due to ↑ release of nitric oxide and cyclooxygenase-dependent relaxing endothelial factors Inhibition of Cytochrome P450 3A4, 1A2, 2C9, 2D6, 3A7, 2A6, 3A5	**Binds to a site on the extracellular aspect of the Na^+/K^+ ATPase pump and inhibits its action** → ↑ in intracellular Na^+ The ↑ Na^+ gets exchanged for Ca^{++} → ↑ intracellular concentrations of Ca^{++} → activation of contractile proteins (actin, myosin) → ↑ inotropy **Direct action on atrioventricular node** → ↓conduction velocity Due to its effects on intracellular calcium concentrations, digoxin may induce apoptosis of tumour cells via a pathway involving mitochondrial cytochrome c and caspases 8 and 3
Formulations	Oral and intravenous Tablets 100 and 200 mg Intravenous 50 mg/ml, 150 mg/100 ml, 360 mg/200 ml	Oral and intravenous Tablets 0.125, 0.25, 0.5 mg Capsules 0.05, 0.1 or 0.2 mg equivalent to 0.0625, 0.125, and 0.25 mg in tablet form. Elixir 0.05 mg/ml. Intravenous 0.1 mg/ml or 0.25 mg/ml
Stability and storage	Store tablets at room temperature Protect from heat, moisture, and light Light protection is not necessary upon administration Amiodarone Hydrochloride 0.6 mg/ml in dextrose 5% in water is stable for five days at room temperature. Solutions containing < 0.6 mg/ml of amiodarone hydrochloride in dextrose 5% in water are unstable and should not be used	Store tablets at room temperature Protect from heat, moisture, and light
Administration	IV: concentrations >2 mg/mL should be administered via a central venous catheter Initial infusion rate should not exceed 30 mg/min Inline 0.22 micron filter should be used	IV doses must be infused over a minimum time period of 5 minutes

	IO: used in emergency situations	
Reconstitution	Dilute with dextrose 5% in water to a concentration of 1 - 6 mg/mL	Dilute with 4 fold or greater volume of 0.9% sodium chloride solution or dextrose 5% in water
Administration incompatibility	Aminophylline, Cefazolin Sodium, Heparin sodium, Sodium bicarbonate, Cefamandole Nafate, Mezlocillin Sodium, Floxacillin, Digoxin	Dobutamine, Doxapram, Amphotericin B Cholesteryl sulphate, Amiodarone, Fluconazole, Foscarnet, Insulin, Propofol
Indication	**IV** - initiation of treatment and prophylaxis of frequently recurring **ventricular fibrillation and hemodynamically unstable ventricular tachycardia in patients refractory to other therapy** **PO** - treatment of life-threatening recurrent ventricular fibrillation and recurrent hemodynamically unstable ventricular tachycardia	Treatment and management of congestive cardiac insufficiency, arrhythmias and heart failure
Therapeutic uses	As above Treatment and recurrence prevention of supraventricular arrhythmias refractory to conventional treatment, especially supraventricular tachycardia associated with Wolff-Parkinson-White syndrome, also, atrial fibrillation, paroxysmal atrial fibrillation, atrial flutter and ectopic atrial tachycardia	As above **Treatment of choice for controlling rapid ventricular rate in patients with atrial fibrillation or flutter** Treatment of established or paroxysmal atrial fibrillation, atrial flutter or paroxysmal atrioventricular junctional rhythm with a fast ventricular rate Treatment of cardiac failure in combination with other agents Prophylactic use to prevent arrhythmias and congestive heart failure in patients with heart disease without failure during certain stressful situations (surgery, severe illness, pregnancy) Termination of pregnancy Treatment of malignancy (experimental)
Dose	Steady-state concentrations of 1 to 2.5 mg/L have been associated with antiarrhythmic effects and acceptable toxicity following chronic oral therapy **PO**: 800-1600 mg daily (divided doses) for 1/52, then 600-800 mg dly for 1-3 weeks, then 100-400 mg dly **IV**: Loading dose - 150 mg over the first 10 minutes (15 mg/min), followed by 360 mg over the next 6 hours (1 mg/min) Maintenance - 540 mg over the remaining 18 hours (0.5 mg/min) During cardiac arrest 300 mg or 5mg/kg (paediatric dose) is used	Dosage should be based on lean or ideal body weight Therapeutic levels are between 0.6 and 2.6 nmol/l and requires monitoring Loading dose PO or IV - administer 50% initially; then may cautiously give 25% 8hly x2 with careful assessment of clinical response and toxicity before each dose **PO**: Loading dose 10-15 mcg/kg Maintenance 3.4-5.1 mcg/kg/day **IV**: Loading: 8-12 mcg/kg total Maintenance: 2.4 to 3.6 mcg/kg/day Loading dose only used in AF, not required in CCF Switching from IV to PO: IV dose (mcg) x 1.25 = PO dose (mcg)
Monitoring	Serum hepatic enzyme concentrations should be monitored at regular intervals Serum electrolyte levels, thyroid, eye and respiratory function periodically	Measure serum digoxin levels 6-10 hours post administration Monitor serum electrolytes and creatinine periodically
Dose modification		Decrease dose in renal impairment
Half life	Elimination $t^{1/2}$ =4h-54 days Biological $t^{1/2}$ = **58 days** (range 15-142 days)	Elimination $t^{1/2}$ = 34-44 hr with normal renal function Biological $t^{1/2}$ = **3.5 - 5 days**
Bioavailability	Slowly and variably absorbed from the GI tract 22 - 86%	Absorption occurs from small intestine ↓ by co-administration of food, malabsorption syndromes, antacids, and ↓ GI motility

		Elixir 70% - 85% Capsules 90% - 100% Tablets 60% - 80%
Protein binding	**>96%**	**25%**
Distribution	**Volume of distribution: adults 1.3-65.8 l/kg** Widely distributed in adipose tissue, liver, lung, spleen, skeletal muscle, bone marrow, adrenal glands, kidneys, pancreas, bile, testes, semen, saliva, lymph nodes, myocardium, thyroid gland, skin, and brain Placental transfer - foetal level 10 - 25% of maternal plasma level Breast milk level substantially higher than maternal plasma level	**Volume of distribution: adults, 7-8 l/kg, neonates 10 l/kg, infants 16 l/kg** Vd ↓ in renal disease, hypothyroidism, quinine therapy and the elderly (↓ muscle mass) 65% of the absorbed dose distributed to skeletal muscle Widely distributed in heart, kidneys, intestine, stomach and liver Lowest concentrations are in plasma, adipose tissue and brain Freely cross the placenta with similar foetal and maternal plasma concentrations Maternal concentrations in plasma and breast milk are similar.
Metabolism and excretion	**Hepatic - extensively metabolised** via CYP2C8, also possible metabolism in the intestinal lumen and/or GI mucosa Mainly biliary excretion Negligible excretion in urine	Hepatic (independent of cytochrome P-450 system) Bacterial metabolism in large intestine **50%-80% of dose excreted unchanged in urine**, 9-13% excreted via faeces and bile
Active metabolites	Major metabolite of amiodarone is Desethylamiodarone (DEA), which also has antiarrhythmic properties and a $t^{1/2}$ = 36 (14-75)	Digoxin undergoes stepwise cleavage of the sugar moieties to form digoxigenin-bisdigitoxoside, digoxigenin-monodigitoxoside, and digoxigenin; with progressively decreasing cardioactivity Other metabolites are cardio-inactive
Clearance	90-158 mL/h/kg (single dose IV of 5 mg/kg over 15 min) 1.9 ml/min/kg (range: 1.4-2.5 ml/min/kg)	Renal clearance 191 ml/min
Contraindications	Known hypersensitivity Cardiogenic shock Sinus node dysfunction associated with sinus bradycardia Second- or third-degree AV block Bradycardia causing syncope, except with functioning artificial pacemaker Porphyria	Known hypersensitivity Ventricular fibrillation
Overdose	Occur with serum level 2.5 mcg/mL	Occur with serum level > 2 ng/mL Toxicity partly due to loss of intracellular potassium **CVS**: Atrial tachycardia, atrial fibrillation, atrial flutter, ventricular tachycardia , ventricular fibrillation, bigeminy, junctional premature complexes, progressive bradyarrhythmias, heart block, heart failure **GI**: Nausea, vomiting, salivation, abdominal pain **CNS**: Headache, dizziness, vertigo, agitation, seizures, visual disturbance, drowsiness, muscle weakness, confusion, coma, respiratory failure
Treatment of overdose	Supportive management, ALS Activated charcoal, cholestyramine Sodium bicarbonate may reverse cardiac depressant effects caused by inhibition of the fast sodium channel Magnesium sulphate for polymorphic ventricular tachycardia Avoid Class 1a agents	**Antidote = Anti-digoxin Fab antitoxin** (for significant toxicity only) Supportive management, ALS Induced vomiting if <30 min since ingestion and patient alert and stable Activated charcoal, cholestyramine Class 1b agents for tachyarrhythmias Atropine or PM for bradyarrhythmias Correct electrolytes

		Avoid calcium
Adverse reactions	**Pulmonary** – occurs in 10-17%, potentially fatal (10%) - interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary fibrosis, exacerbation of asthma, haemoptysis **Neurological** - peripheral neuropathy, proximal muscle weakness, visual disturbances in 10%, asymptomatic corneal deposits in almost all patients, rarely reports of nystagmus, ischemic optic neuritis, optic neuropathy, papilledema, corneal degeneration, scotoma, lens opacities, ocular discomfort and macular degeneration **Endocrine** – hypothyroidism in 2-4%, rarely hyperthyroidism, abnormalities of liver function test results in 3-55%, rarely, severe hepatic injury (hepatitis, hepatocellular necrosis, cirrhosis), non-infectious epididymitis or epididymo-orchitis and/or scrotal pain, gynecomastia, hyper/hypoglycaemia **CVS** – arrhythmias (brady and tachy) in 2-5%, hypotension in 16% (IV), new or worsening cardiac failure in 2-3%, phlebitis **Coagulation** - abnormalities in 1-3%, rarely severe thrombocytopenia **GI** - nausea, vomiting, constipation and anorexia in 25%, abdominal pain, abnormal salivation and abnormal taste in 1-3% **Dermatological** – photosensitivity in 10%, pigment deposition → blue-gray discoloration of the skin, rash and hair loss in <1, rarely toxic epidermal necrolysis and psoriasis Stevens-Johnson syndrome in <2% of patients receiving the drug IV **Malignancy** - possible association between amiodarone and an increased risk of cancer, especially in males	**GI** – anorexia, nausea and vomiting, abdominal pain, diarrhoea, constipation in 1-10%, very rare - intestinal ischemia, intestinal haemorrhagic necrosis **Neurological** - visual disturbances, headache, facial pain, weakness in 1-10% **Psychiatric** – depression, apathy, anxiety, psychosis, confusion, delirium, hallucination in 0.1-1% **Endocrine** – gynecomastia **CVS** - arrhythmias (brady and tachy) in 1-10%, **Dermatological** – Rashes in 1-10%
Hypersensitivity	Rare Angioedema and anaphylactic shock have been reported	Rare Usually within 6-10 days after initiating therapy. Rashes, usually accompanied by eosinophilia Urticaria, pruritis, fever Facial, angioneurotic, or laryngeal oedema; Alopecia of the scalp, desquamation Shedding of finger and toe nails Rarely, thrombocytopenic purpura
Drug interactions	↑ **levels of other agents** e.g. Digoxin, Phenytoin, Procainamide, Flecainide, Cyclosporine, Quinidine, Theophylline, Sildenafil Ledipasvir, Sofosbuvir – severe bradycardia Warfarin – serious ↑ in PT Simvastatin - ↑risk of rhabdomyolysis Potentially serious adverse	↑ **levels**: Amiodarone, Quinidine, Verapamil, Nifedipine, Diltiazem, Nicardipine, Flecainide, Propafenone Prazocin , Calcium IV, Spironolactone, Furosemide, Indomethacin, short course Ibuprofen Diazepam, Heparin, Methyldopa, Trazodone, Tolbutimide, Erythromycin, Tetracycline, Metoclopramide, Hydroxychloroquine, Cyclosporine, Trimethoprim

	cardiovascular effects have occurred in some amiodarone treated patients undergoing general anaesthesia, e.g bradycardia and heart block resistant to atropine and adrenaline, these patients may require temporary pacing perioperatively ↓**levels**: Activated charcoal, Cholestyramine	↓**levels**: Sulfasalazine, Neomycin, Penicillamine, Aminosalicylic acid, Activated charcoal, Cholestyramine
Food interactions	Grapefruit juice significantly ↑ levels	↓ absorption with co-administration of food Hawthorn ingestion increases risk of arrhythmias
Special precautions	Avoid during pregnancy and breastfeeding Use with caution in patients receiving calcium channel blockers and/or beta-adrenergic blockers Use with caution in patients receiving agents known to prolong QT interval IV administration not recommended for use in the paediatric population due to potential for adverse male reproductive tract development and the preservative benzyl alcohol has been associated with the potentially fatal "gasping syndrome" in neonates	Use with extreme caution during pregnancy and breastfeeding Use with extreme caution, if at all, in patients with idiopathic hypertrophic subaortic stenosis because increased obstruction to left ventricular outflow may result Incidence of toxicity increased by hypokalaemia, hypomagnesaemia, hyperkalaemia, hypernatraemia and alkalosis Risk factors for toxicity include hypothyroidism, renal failure, myocardial infarction, myocarditis, chronic constrictive pericarditis, hypoxia, severe pulmonary disease, recent cardiac surgery, severe bradycardia, severe heart failure, sick sinus syndrome, ventricular tachycardia, ventricular premature contractions, Wolff-Parkinson-White syndrome and reduced muscle mass It should be avoided or used very cautiously in these patients

References:

1. Santana, Luis F.; Cheng, Edward P.; Lederer, W. Jonathan, How does the shape of the cardiac action potential control calcium signaling and contraction in the heart? Journal of Molecular and Cellular Cardiology: 2010; 49 (6): 901–903.
2. T. E. Peck, S. A. Hill & M. Williams Pharmacology for Anaesthesia and Intensive Care (3rd edition) Cambridge University Press
3. Ziff, Oliver J; Lane, Deirdre A; Samra, Monica; Griffith, Michael; Kirchhof, Paulus; Lip, Gregory Y H; Steeds, Richard P; Townend, Jonathan; Kotecha, Dipak, Safety and efficacy of digoxin: systematic review and meta-analysis of observational and controlled trial data. BMJ: 2015; h4451
4. Tagami T., Yasunaga H, Yokota H, Antiarrhythmic drugs for out-of-hospital cardiac arrest with refractory ventricular fibrillation. Critical Care: 2017; 21, 59.
5. Park H, Kim Y, Adverse effects of long-term amiodarone therapy, Korean J Intern Med 2014;29:571-573
6. WHO Model List of Essential Medicines (19th List) (PDF). World Health Organization. April 2015
7. Databases:
   http://www. pubchem.ncbi.nlm.nih.gov
   http://www.clinicaltrials.gov
   http://www.cochrane.org

Altitude and Anaesthesia

Prof. MFM James

UCT Dept of Anaesthesia & Perioperative Medicine
Emeritus Professor

Inhalational pharmacology is the backbone of clinical anaesthetic practice. The delivery of respiratory gases, together with inhaled anaesthetic agents, is the sine qua non of anaesthesia. An understanding of how gases behave under conditions of altered barometric pressure is, therefore, essential for all anaesthetists, particularly those practising in a country like South Africa where there are wide variations in altitude between major cities.

Oxygen and Altitude

While the proportional composition of the atmosphere remains remarkably consistent, ambient pressure decreases logarithmically with ascent, causing a corresponding decrease in the partial pressure of all atmospheric gases, but, most importantly, that of oxygen (PO_2). The physiological effects of altitude are predominantly due to the resultant hypoxia and hypobaria. High altitude is defined as >1 500 m above sea level, where the physiological effects of altitude may first be consistently observed, but pathological consequences are very rare. At this level, the alveolar partial pressure of oxygen (PaO_2) is ~10 kPa, compared with the sea level value of 13 kPa(Figure 1) [1]. It is worth noting that large areas of SA, including the extensive metropolitan areas in Gauteng, therefore qualify as high altitude.

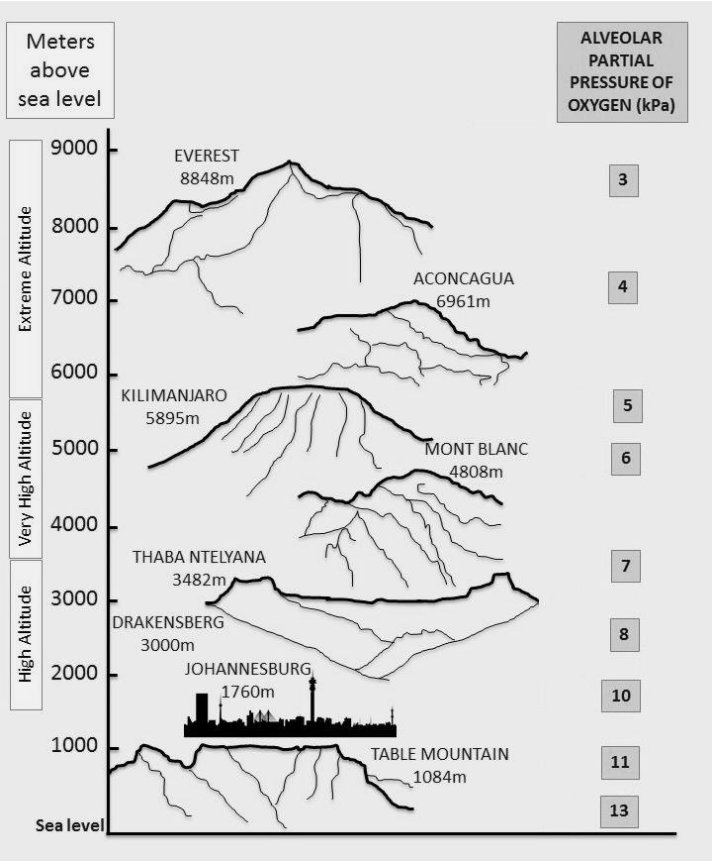


Figure 1. Changes Alveolar PO2 with altitude. Adapted with permission from reference 1.

The partial pressure of oxygen in the lung is given by the alveolar gas equation:

$$P_A0_2 = F_i0_2(P_B\text{-}PH_20) - PaCO_2/R$$

At sea level	$P_A0_2 = 0.21(101\text{-}6.2)\text{-}5.3/0.8$	=	13.3 kPa
In Johannesburg	$P_A0_2 = 0.21(80\text{-}6.2)\text{-}4.3/0.8$	=	10.13 kPa
Summit of Everest	$P_A0_2 = 0.21(34\text{-}6.2)\text{-}1.5/0.8$	=	3.9 kPa

Gas volumes are absolutely dependent on the barometric pressure to which the gasses are exposed (Boyle's Law) and thus fractional concentration does not adequately express the active component of any gas within a mixture. For example, 21% oxygen at sea level represents a partial pressure of 21 kPa at sea level. However, the same fractional concentration at the top of Mount Everest represents a partial pressure only approximately 7 kPa (see Figure 1). Since physiology entirely depends on the solution of these gases in the tissues, which is partial pressure dependent, the use of fractional concentration or percentage is profoundly misleading. Thus a person living in Johannesburg has a persistently lower P_A02 than the same individual at sea level and a mountaineer on Everest is profoundly hypoxic – all at 21% oxygen!

To understand the physiological impact of this issue properly, it is necessary to revisit some basic laws of physics. Dalton's law of partial pressure states that the partial pressure (PP) of one gas in a mixture of gases is the pressure that gas would exert if it alone occupied the whole space. Thus if all gasses other than oxygen are removed from a container of air at sea level while keeping the volume constant, the remaining gas would be 100% oxygen, but the pressure in the container would be only 21 kPa. The second important physical principle is expressed in Henry's Law which states that, at constant temperature, the amount of a given gas that dissolves in a given liquid is directly proportional to the partial pressure of that gas in equilibrium with that liquid. The implication of this is that the number of gas molecules entering the blood stream from the alveolus is dependent on the PP (not the concentration) of that gas in the alveolus. Finally, Ostwald's solubility coefficient is the volume of gas dissolved per milliliter of liquid and per unit PP of the gas at any given temperature, with the solubility increasing as the liquid temperature decreases. The corollary of this is that the tendency of a gas to leave a solution is also driven by its partial pressure and the temperature of the liquid. This is why a carbonated drink bubbles when opened as the CO_2, dissolved under pressure, escapes from the liquid when that pressure is released (especially if the liquid is warmed).

The gradient between PaO_2 and the oxygen PP in the tissues provides the driving force for oxygen to reach the mitochondria resulting in the oxygen cascade (Figure 2).

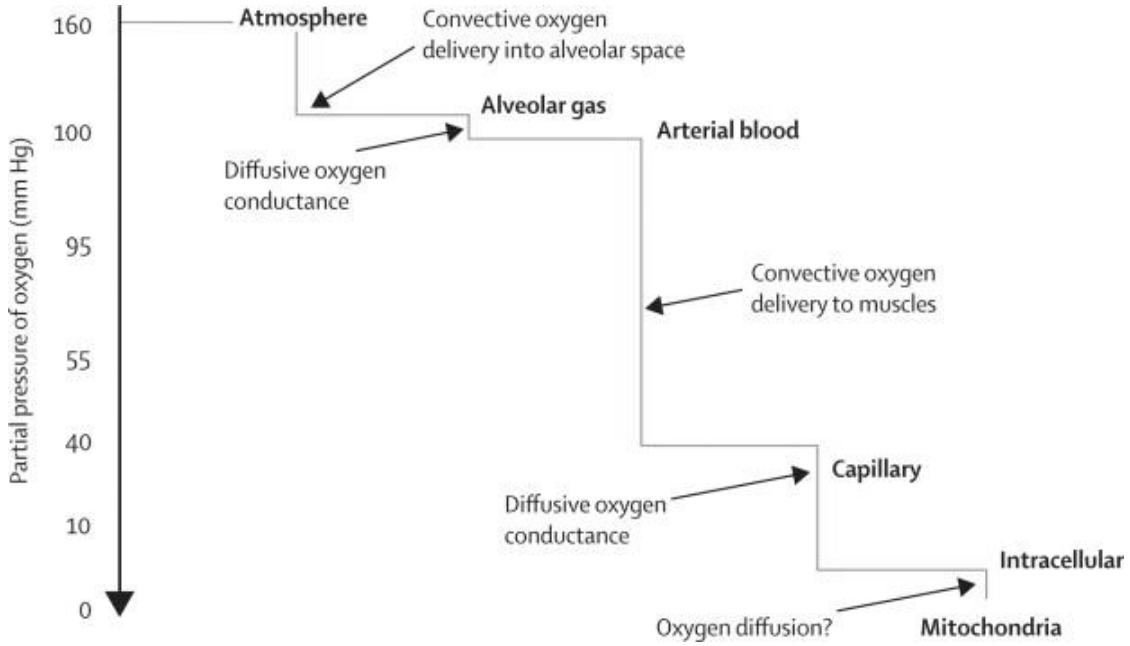


Figure 2. Oxygen cascade illustrating the partial pressure gradients that drive oxygen across the various tissue components to reach the mitochondria.

A quarter of the world's population live at more than 500m above sea level and nearly 10% reside over 1500m [2] including heavily populated cities such as Denver, Nairobi, Bogota, Mexico City and Johannesburg. Inhalational anaesthesia at such altitudes is common. South Africa is a country with widely varying altitudes and therefore it is imperative that anaesthetists understand the influence of altitude on the accuracy and reliability of their delivery and measurement systems.

Inhalational Agents

Anaesthetic gases and vapours exert their effects in proportion to the partial pressure that can be established within the brain. The driving force for an anaesthetic agent to enter the brain from the blood is the partial pressure of the anaesthetic in the blood, not its concentration. This explains the apparent anomaly of agents such as ether, which are very soluble in blood, and thus develop a high concentration rapidly, but act very slowly. The high solubility of ether means that the partial pressure in the alveolus falls rapidly, resulting in a low partial pressure in the alveolus and thus in the blood. The driving force for the agent to enter the brain is thus low until such time as the blood is saturated with ether and the partial pressure begins to rise. The same phenomenon explains why relatively insoluble agents such as desflurane exert their effects rapidly, but require a high alveolar partial pressure to dissolve sufficient gas molecules in the blood to exert the pharmacological effect. The reverse is true when it comes to removal of the gases, meaning that insoluble gases with high partial pressures have a high driving force to leave the blood when the gas is removed from the alveolus resulting in more rapid recovery than that seen with soluble agents.

The potency of an anaesthetic agent is classically described in terms of the minimal alveolar concentration (MAC) defined as the concentration of an inhaled agent that will prevent movement to a surgical stimulus in 50% of subjects. Note that concentration is specified; this means that, for precision, the atmospheric pressure should be quoted when stating a MAC value. If the pressure departs significantly from 1 atm, then this becomes essential. The difficulty is avoided by expressing MAC as a partial pressure (MAP).[3] At sea level, 1 MAC of isoflurane is 1.15 vol %. This corresponds to an end-tidal partial pressure of 101.3 kPa X 0.0115 = 1.16 kPa. To achieve an end-tidal isoflurane partial pressure of 1.16 kPa in Mexico City (barometric pressure 76.4 kPa), one should aim for an end-tidal concentration of 1.16 kPa/76.4 kPa = 1.5%. Consequently, when using isoflurane in Mexico City, its MAC is 1.5 vol%[4], not the standard 1.15%. It is far more logical, therefore, to use the concept that is independent of altitude, and thus applicable throughout the world, minimal alveolar partial pressure (MAPP). In SI unit terms this conversion is simple as the kPa MAPP value and the sea level MAC value are essentially the same [5]. The correct measure of potency that is independent of altitude is the minimum alveolar partial pressure (MAPP), first described in 1984[6] and now accepted as standard in the major anaesthetic textbooks.[7]

Nitrous oxide, although a vapour at normal room temperature, behaves like a gas with a MAPP of 105 kPA (about 105% MAC at sea level). As it behaves like a gas, not a vapour, the partial pressure of any given concentration decreases as barometric pressure falls. Consequently, the value of nitrous oxide as an anaesthetic agent diminishes with increasing altitude. At sea level, N_2O has a considerable analgesic potential, approximately equivalent to 10 mg IV morphine. However, at the altitude of Johannesburg, the 20% reduction in partial pressure of the agent reduces its analgesic effect by half and at 3 000 m there is no detectable analgesic effect at all (Figure 3).[8]

These results have been widely confirmed. The reduced effectiveness of N2O at the altitude of Johannesburg was noted as long ago as 1913 [9]. Powell[10] concluded that the efficacy of N2O was decreased at an altitude of 5280 feet (1670 m). Cleaton-Jones and colleagues[11] reported that there were only marginal differences between a group studied at sea level and a different group studied at 1670 m. However, they did not measure objective end points and there were strong trends suggesting a lower potency of N2O at altitude, particularly in conscious level, at 60 and 70% concentrations between the two altitudes. A study of anaesthesia using volatile agents and N_2O in combination at sea level and at altitudes >1300 m showed a significantly greater consumption of volatile agents at higher altitudes, suggesting a diminished effect of N2O [12]. Furthermore, a comparative trial of i.v. anaesthesia in conjunction with 66% N_2O at high and low altitudes demonstrated significantly higher propofol requirements at altitude [13]. It can be safely concluded that the efficacy of N_2O is entirely dependent on the partial pressure of the agent and not the concentration.

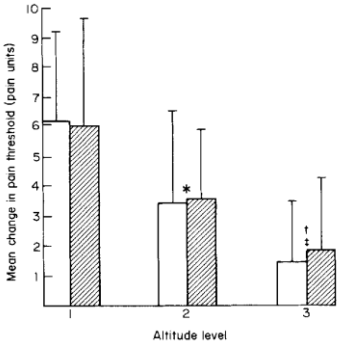


Figure 3. Effect of 50% nitrous oxide on the analgesic threshold in volunteers at sea level (level 1), 1 700 m (level 2, equivalent to Johanesburg) and at 3 000 m (level 3) in either pure oxygen (open bars) or an air/oxygen mixture (hatched bars). At 3 000 m there was no detectable analgesic effect.[8]

Vaporisers

Variable bypass anaesthetic vaporisers function on the basis of diluting a saturated vapour of the agent into a bypass flow stream. The mass of anaesthetic agent delivered from the vaporiser is determined by the SVP of the agent within the vaporising chamber at a fixed temperature and a variable splitting ratio, which dilutes the SVP to the desired value. Since SVP is solely determined by temperature and virtually unaffected by barometric pressure, temperature compensation measures are used within the vaporiser to minimise temperature fluctuations and to adjust the splitting ratio should the temperature vary from the calibrated value (usually 20°C). These vaporisers deliver a partial pressure, not a concentration of the agent and continue to be reasonably accurate with fluctuations in barometric pressure.[7,14] Thus, a variable bypass vaporiser set to deliver 2% at sea level will deliver a partial pressure of approximately 2 kPa regardless of the barometric pressure. The same vaporizer in Johannesburg or Denver, Colorado (at an altitude of approximately 1700m (5500 feet) and a barometric pressure of approximately 80 kPa will continue to deliver the same partial pressure of 2 kPa, but at a concentration of 2.5%. At a higher altitude where the barometric pressure is half that at sea level, the proportion of isoflurane vapour output increases due to the lower barometric pressure. Therefore, the dial setting that delivered 2% isoflurane at sea level will now deliver 4% isoflurane. However, according to Dalton's law, the partial pressure of isoflurane delivered would be approximately the same at both altitudes since 2% isoflurane at 760mm Hg (15.2 mm Hg) is the same as 4% isoflurane at 380mm Hg (15.2 mm Hg).[1] If the anaesthetist were to titrate anaesthetic agent dosage to end-tidal volatile concentration, the partial pressure would be half that delivered at sea level and the risk of under-dosing (and thus accidental awareness under anaesthesia) would be substantial. The desflurane vaporiser functions on a different principle, more akin to a nitrous oxide cylinder, and delivers a concentration of vapour which requires adjusting upwards for increasing altitude.[15] Cassette-type vaporisers for most volatiles behave like a computer-controlled variable bypass vaporiser. However, the cassette for desflurane injects a known mass of volatile into the system (similar to a carburettor) and thus should also deliver a partial pressure, not a concentration, so similar considerations could apply although this has never been scientifically tested.

Gas Analysers

In gas analysis, volumetric measurement was historically the standard technique used for establishing the fractional concentration of different gases in a mixture using such equipment as the Haldane and Van Slyke devices. Hence the universal use of concentration as the standard descriptor of gas mixtures. These are clumsy, time-consuming techniques that are not amenable to the rapid gas analysis required in current clinical practice. Modern gas analytical techniques all measure gas composition using a specific property of each gas, ignoring other gases in the mixture.

It is obviously vital to know and understand how we measure the delivery of oxygen and removal of carbon dioxide as errors in this area may be life-threatening. In addition, anaesthetic vapour analysis is an essential element of monitoring during inhalational general anaesthesia [16] and recent publications have re-emphasised the importance of monitoring end-tidal anaesthetic concentrations as a means of minimising the risks of awareness in this setting [17,18]. Consequently the question arises as to whether exhaled vapour concentration is a valid measure of the dose of volatile anaesthetic gases in all circumstances [5]. In particular, are such measures misleading under conditions of altered barometric pressure occurring with increasing altitude?

Oxygen Analysers

There are four main types of analysers currently in common use for measurement of the oxygen content of gas mixtures. They are paramagnetic, fuel cell, oxygen electrode, and mass spectrometer devices. All respond to the partial pressure of oxygen alone, and the output changes as the barometric pressure changes. At an altitude of 5000 ft an oxygen analyzer set to measure 21% oxygen at sea level will give a reading of 17.4% unless it is recalibrated to read 21% in air at the new pressure (Table 1)[5].

[1] In actual fact, the splitting ratios vary slightly with altitude, so variable bypass vaporisers at altitude actually deliver a slightly higher partial pressure than that marked on the dial, but the error is much smaller (and in a safer direction) than the concentration error.

TABLE 1. INFLUENCE OF BAROMETRIC PRESSURE ON OXYGEN CONTENT OF AIR AS MEASURED BY PARAMAGNETIC OXYGEN ANALYSER

	Barometric Pressure			Partial Pressure of Oxygen	
Altitude (ft)	(kPa)	(mm Hg)	Scale Reading (% oxygen)	(kPa)	(mm Hg)
0	101.5	760	21.0	20.9	165
5000	83.2	624	17.4	17.4	122
10000	69.0	518	14.2	14.2	110

If the device were to be calibrated in terms of partial pressure the scale readings would reflect the true state of oxygen availability to the patient. This effect is of relatively minor importance unless accurate research work is being performed or the changes in barometric pressure are extreme. At altitudes of 10,000 ft and above, 21% oxygen becomes a relatively hypoxic mixture, and measurements of partial pressure are the only sensible ones to follow. It is worth noting that breathing air at this altitude is equivalent to breathing 14% oxygen at sea level. As a practical point, the gain controls on some models of oxygen analysers do not permit resetting of the scale to read 21% in air at such low partial pressures of oxygen. These devices will, however, continue to reflect the reduced partial pressure. Under hyperbaric conditions, it is again important to know the partial pressure of oxygen in the mixture inhaled, because 21% oxygen at 5 atmospheres pressure will exert a partial pressure greater than that of 100% oxygen at sea level, with the consequent risks of oxygen toxicity. Hence, divers operating at depths of 200 metres (a pressure of +20 atm), such as in the North Sea oil fields use gas mixtures of 1% oxygen in helium, which results in a partial pressure of 20 kPa of oxygen under 20 atm pressure.

Carbon Dioxide Analysers

Carbon dioxide is most commonly measured by absorption of infrared radiation by the gas. Such instruments are often produced commercially with an internal calibration device that is supposed to read in percentages. Because the instrument actually responds to partial pressure, this is misleading, and reliance on the internal calibration device will introduce errors of clinical importance when such equipment is used at altitude. A device with an internal calibration nominally set at 7% will, in fact, be calibrated to 7 kPa (53 mmHg), which is actually 8.2% of barometric pressure at 5000 ft altitude. For such equipment to function correctly, it must be either recalibrated against known concentrations of carbon dioxide at the correct barometric pressure or the scale converted to read partial pressure. If kilopascals are used, this conversion is simple because sea level percentages and kilopascals are, for practical purposes, the same. Fortunately, many CO_2 analysers now available display the transducer output in pressure units and not as percentages and these should be used wherever possible. The effect of varying barometric pressures on the output of an infrared analyzer using a fixed concentration of carbon dioxide is shown in Table 2. At an altitude of 5000 ft, the concentration of carbon dioxide in alveolar gas is 5.5%. If a device calibrated in percentages is used to monitor end tidal CO_2; to maintain normocapnoea, the use of the percentage figure will result in the patient being rendered significantly hypercapnic, because the partial pressure of expired CO_2 will be 5.5 kPa (42 mmHg), the normal at this altitude being 4.6 kPa (35mmHg). A more subtle error may also be introduced if a machine of this nature is used to monitor a patient being ventilated inside a one-man hyperbaric chamber in, for example, the management of carbon monoxide poisoning. If the analyser is placed outside the chamber and the patient's expired gas led outside the chamber before measurement, the gas will expand. The concentration of carbon dioxide in the mixture will remain unchanged, but the partial pressure will decrease, thus giving a falsely low reading.

TABLE 2. INFLUENCE OF BAROMETRIC PRESSURE ON 4.5% MIXTURE OF CO_2 AS MEASURED BY INFRARED ANALYSER

	Barometric Pressure			Partial Pressure of Oxygen	
Altitude (ft)	(kPa)	(mm Hg)	Scale Reading (% CO_2)	(kPa)	(mm Hg)
0	101.5	760	4.5	4.5	34.2
5000	83.2	624	3.7	3.74	28.5
10000	69.0	518	3.1	3.18	24.3

The same principle applies to all modern gas measurement devices including those measuring oxygen and anaesthetic agents (Table 2).

Vapour Analysers
Similar arguments apply to the use of vapour analysers, all of which in modern practice respond to partial pressure, although they are almost invariably calibrated in percentages (the latest Dräger analyser can be set to return measurements in partial pressure). These devices are almost all based on infrared absorption technology (as is used in capnography) and thus the same considerations apply. This situation can be confirmed by producing a gas mixture of known composition by vapourising a known mass of a volatile anaesthetic liquid into a closed vessel of known volume. Reduction of the ambient pressure to which the flask and its contents are subjected will result in a fall in the partial pressure of the vapour in the flask, with a consequent reduction in the reading of the analyzer, even though the concentration of the vapour is unchanged. Alternatively, a fresh vapour mixture can be prepared at different ambient pressures by adding the same mass of liquid to the cleaned flask after the pressure had been altered. In this way, a constant mass of anaesthetic vapour would be contained within the flask and the partial pressure of the vapour should be the same at each level of pressure, although the concentration would be different. The output of the analyzer remains constant despite the altered concentration.

The standard device used to calibrate the output of vaporisers is the Rayleigh refractometer. This instrument compares the refractive index of the gas under investigation with that of a standard gas mixture and produces an accurate measure of the number of gas molecules in the sample. Again such devices respond to partial pressure not concentration, meaning that vaporiser output is calibrated in terms of partial pressure (even though expressed as a concentration).

As with other gas analysers, the use of percentage measures on an agent analyser to guide the use of anaesthetic vapours at altitude can lead to serious dosing errors. It must be remembered that vapour analysers measure partial pressure not concentration. The use of percentage units to describe anaesthetic agent dosing is simply wrong and can lead to serious errors in dosages, particularly at altitude.

Gas density and flow

Changes in barometric pressure produce changes in the density of gases. As a result, an apparatus whose function is partly or wholly dependent on gas density will not behave as it would at the altitude at which the device was calibrated (almost invariably sea level). Consequently, equipment used by anaesthesiologists in the operating room, critical care area, and laboratory may not function in the expected manner.

Flowmeters
Most flowmeters use the decrease in pressure that occurs when a gas passes through a resistance as a measure of gas flow. The magnitude of the decrease in pressure depends on the density and viscosity of the gas. In situations where the resistance represents an orifice, resistance depends primarily on the density of the gas. Where the resistance is tubular, viscosity becomes the prime determinant of the magnitude of the decrease in pressure provided that the flow remains laminar. Most flowmeters use a floating ball or bobbin supported by the stream of gas in a tapered tube. At low flow rates, the device depends primarily on tubular flow, and as the float moves up the tube the resistance behaves progressively more like an orifice. The density of a gas changes, of course, with changes in barometric pressure, but the viscosity changes relatively little, as it is primarily dependent on temperature. Gas flow through an orifice is inversely proportional to the square root of the density of the gas. As the density of the gas falls, therefore, the flow through an orifice of given size will increase. Thus at altitude the actual flow delivered by a flowmeter will be greater than that indicated by flowmeter. The actual flow delivered by a flowmeter under conditions of altered barometric pressure can be described by Equation 1:

$$\text{Equation 1: } F_1 = F_0 \sqrt{\frac{\rho_0}{\rho_1}}$$

where F_1 is the flow delivered at the new pressure,
Fo is the flow delivered at the original pressure,
ρ_0 is the original density of the gas,
ρ_1 is the density of the gas at the new pressure.

As density is directly proportional to pressure, values for barometric pressure may be substituted for density. The relative contributions of density and viscosity to the behaviour of a floating bobbin flowmeter are unpredictable, as each flowmeter has its own characteristics determined by the relationship between the shape of the bobbin and the taper of the tube. Thus, the exact role of viscosity in the determination of the position of the float has not been well established. Viscosity has been shown to exert an important effect in clinically used flows, and at flow rates at which viscosity is the predominant factor, Equation 1 will not apply. Little if any error in measured flow will occur with changes in altitude at the lower settings of the flowmeter where viscosity is the main determinant of the pressure drop.

If accurate measurements are to be made, the only practical approach is to recalibrate the flowmeter at the altitude at which it is to be used. These comments do not apply to fixed orifice flowmeters, in which turbulent flow is probable at all flow rates, and the error should be accurately described by Equation 1 at any flow. Various other flow-measuring devices will also perform inaccurately at altitude. The manufacturers of the Wright respirometer include a guide to the inaccuracy of the instrument at altitude. Devices that depend on other physical properties of the gas, such as katharometers and electronic flowmeters, will also tend to under-read, because the number of gas molecules per unit volume, and hence the thermal or electrical capacity, will be reduced. Pneumotachographs, on the other hand, use laminar flow to generate a pressure drop, and therefore viscosity, not density, will be the prime determinant of the measurement obtained. These devices should, therefore, continue to perform well regardless of changes in barometric pressure but may be sensitive to temperature fluctuations.

High Air Flow Oxygen Enrichment Devices
The importance of administering known concentrations of oxygen at flows exceeding peak inspiratory flow in the management of patients with respiratory disease has often been stressed. High airflow oxygen enrichment equipment usually consist of a fixed orifice Venturi device for which a specified minimum gas flow must be provided in order to produce adequate flow rates. Changes in barometric pressure might be expected to exert profound effects on such devices, because the driving force that accelerates air along the breathing pathway is the pressure gradient between the atmosphere and the negative pressure area created by the jet of oxygen emerging from the nozzle. Under conditions of reduced barometric pressure, these devices might be expected to "run rich," that is, to produce higher oxygen concentrations than those set but at a lower flow rate than that delivered at sea level. When tested, however, these devices performed better than might have been anticipated. In virtually every case, there was a small but consistent increase in the oxygen percentage and a similar decrease in the total flow produced by the device. The magnitude of these errors, however, was smaller than had been anticipated when a standard flowmeter, not corrected for altitude, was used. When a fresh gas flow is set on a standard flowmeter, it appears that the increased flow delivered by the flowmeter at altitude largely offsets the errors in delivered flow and oxygen concentration that would otherwise occur with decreases in barometric pressure. When flows corrected for altitude are used, the performance of the device deteriorates to the point where the flow delivered by the mask may well decrease below the patient's peak inspiratory flow, and higher than expected percentages of oxygen are produced. This is unlikely to be of major importance at altitudes of 5000 ft or less but may be problematic at higher altitudes. It should be noted that although the oxygen percentage may have remained more or less constant, there is a reduction in its partial pressure, and this effect must be allowed for when prescribing such a device for patient use.

Minimal Alveolar Partial Pressure

These considerations have a major import for the outdated concept of minimum alveolar concentration (MAC).

Since all gaseous agents exert their physiological effects in proportion to their partial pressure, the instruments that measure the amount of gaseous agent in any gas mixture measure partial pressure and anaesthetic vaporises deliver only partial pressures of the agents, it is quite simply absurd to persist with percentage expressions of these agents. Using percentages may lead to serious clinical errors. Particularly at altitude the use of percentages as a descriptor of the dose of gases in the clinical situation will inevitably lead to serious clinical errors including possible hypoxia and awareness. This inappropriate terminology should be abandoned. This is clearly illustrated by the so-called “universal definition of ARDS”: P_aO_2/FiO_2 [19]. This is anything but universal as alveolar oxygen PP will be substantially lower at altitude for the same FiO_2 leading to many more patients being wrongly diagnosed with ARDS at altitude.

The only rational solution is to express all medical units of measurement in terms that actually describe the clinically relevant amount. Thus gases should be described in terms of partial pressures and drugs in solution should only be described in terms of the mass of drug per unit of solvent. Anything else will inevitably result in errors in drug dosage and the literature is replete with descriptions of such errors with profound clinical consequences.

Partial pressure is the factor determining the effectiveness of the volatile agents as well as of the inhaled gases. Consequently, because the concentration of an agent required to produce a given effect increases with reductions in barometric pressure, the concept of MAC does not apply accurately at altitude and should be converted to minimal alveolar partial pressure (MAPP). The idea has much to recommend it; using this concept would eliminate many of the problems described herein. For ease of reference, the MAC and MAPP values of the commonly used anaesthetic agents are listed in Table 5.

TABLE 5. EQUIVALENT VALUES OF MAC AND MAPP AT VARIOUS BAROMETRIC PRESSURES

	MAC			**MAPP**	
Agent	Sea Level	5000 ft	10 000ft	kPa	mm Hg
Nitrous Oxide	101.5	126.5	152.2	106.1	798
Halothane	0.75	0.9	1.09	0.76	5.7
Isoflurane	1.2	1.45	1.73	1.22	9.1
Sevoflurane	2.0	2.4	3.2	2.12	15.9
	6.0	7.25	9.6	6.36	47.7

References

1. Hofmeyr R, Meyer W, James M, De Decker R. Recognising and mitigating the risk of altitude-related illness. S Afr Med J 2016;106:834-5.
2. Cohen JE, Small C. Hypsographic demography: the distribution of human population by altitude. Proceedings of the National Academy of Sciences of the United States of America 1998;95:14009-14.
3. White D. Uses of MAC. British journal of anaesthesia 2003;91:167-9.
4. Gorelov V. Calibration of vaporisers. Anaesthesia 2005;60:420.
5. James M, Hofmeyr R, Grocott M. Losing concentration: time for a new MAPP? British journal of anaesthesia 2015:aev151.
6. James MF, White JF. Anesthetic considerations at moderate altitude. Anesthesia and analgesia 1984;63:1097-105.
7. Venticinque SG, Andrews JJ. Inhaled Anesthetics: Delivery Systems Miller's Anesthesia: Elsevier Health Sciences, 2015:2674-704.
8. James MF, Manson ED, Dennett JE. Nitrous oxide analgesia and altitude. Anaesthesia 1982;37:285-8.
9. Firth PG, Pattinson KT. Anaesthesia and high altitude: a history. Anaesthesia 2008;63:662-70.
10. Powell JN. Some aspects of nitrous oxide anesthesia at an altitude of one mile. Anesthesia and analgesia 1969;48:680-5.
11. Cleaton-Jones P, Moyes DG, Whittaker AM. Clinical effects of nitrous oxide and oxygen mixtures at sea-level and at 1700 metres altitude. Anaesthesia 1979;34:859-62.
12. Coetzee JF, Stewart LJ. Fresh gas flow is not the only determinant of volatile agent consumption: a multi-centre study of low-flow anaesthesia. British journal of anaesthesia 2002;88:46-55.
13. Puri GD, Jayant A, Dorje M, Tashi M. Propofol-fentanyl anaesthesia at high altitude: anaesthetic requirements and haemodynamic variations when compared with anaesthesia at low altitude. Acta Anaesthesiol Scand 2008;52:427-31.
14. Eales M, Cooper R. Principles of anaesthetic vaporizers. Anaesthesia & Intensive Care Medicine 2007;8:111-5.
15. Andrews JJ, Johnston RV, Jr. The new Tec6 desflurane vaporizer. Anesthesia and analgesia 1993;76:1338-41.
16. Checketts M, Alladi R, Ferguson K, Gemmell L, Handy J, Klein A, Love N, Misra U, Morris C, Nathanson M. Recommendations for standards of monitoring during anaesthesia and recovery 2015: Association of Anaesthetists of Great Britain and Ireland. Anaesthesia 2016;71:85-93.
17. Pryor KO, Hemmings HC, Jr. NAP5: intraoperative awareness detected, and undetected. British journal of anaesthesia 2014;113:530-3.
18. Punjasawadwong Y, Phongchiewboon A, Bunchungmongkol N. Bispectral index for improving anaesthetic delivery and postoperative recovery. The Cochrane Library 2007.
19. Villar J, Perez-Mendez L, Blanco J, Anon JM, Blanch L, Belda J, Santos-Bouza A, Fernandez RL, Kacmarek RM, Spanish Initiative for Epidemiology S, Therapies for AN. A universal definition of ARDS: the PaO2/FiO2 ratio under a standard ventilatory setting--a prospective, multicenter validation study. Intensive Care Med 2013;39:583-92.

Breathing System Filters

Professor Andrew Levin

Executive Head
Department of Anaesthesiology and Critical Care
University of Stellenbosch

Notes

Notes

Receptor Physiology

Dr Kerry Timmerman

UCT Dept of Anaesthesia and Perioperative Medicine

Receptors are proteins. According to Dorland's Medical Dictionary, a receptor may be defined as "a molecular structure within a cell or on the surface characterized by (1) selective binding of a specific substance and (2) a specific physiologic effect that accompanies the binding".

Receptors may be located at various sites:

- Incorporated into or associated with the cell membrane (generally for ligands that do not penetrate the cell); and
- Intracellular sites (generally for lipid-soluble ligands that can diffuse through the cell wall, or for intermediary messengers generated within the cell)
  - membranes of intracellular organelles
  - cytosol
  - nucleus

Ligands may be natural or synthetic, and can be grouped into different efficacy classes:

- *full agonists* are capable of maximal receptor stimulation
- *partial agonists* are unable to elicit full activity even at saturating concentrations
- *antagonists* have no effect on signaling activity but can prevent other ligands from binding to the receptor
- *inverse agonists* reduce the level of basal or constitutive activity below that of the unbound receptor (i.e. are able to exert an effect opposite to that of the ligand)

The interaction of ligands with receptors is an essential element of cellular communication and key in bringing about changes in cell function. As anaesthetists, we use drugs on a daily basis that interact with receptors, and an understanding of receptor physiology is essential knowledge.
It is beyond the scope of this talk to mention all the currently known receptors, but I have tried to keep it relevant to anaesthesia by giving some clinical examples in each broad group.

Classification of receptors:

1. Ion transport proteins
2. Metabotropic or G-protein coupled receptors
3. Catalytic receptors
4. Intracellular receptors

1. **Ion transport proteins** can be broadly divided into 2 classes, the ion channels and carrier-type transporters, most (but not all) of which possess several transmembrane-spanning domains that create a pore through which ions pass.

   a. Ion channels

   Generally, ion channels facilitate rapid trans-membrane translocation of ions down their concentration and electrical gradients with little or no energy expenditure, i.e. ion movement is passive and usually fast over a short time period. Whether the channel is open ("gated") or closed depends on extrinsic factors such as changes in membrane potential (voltage-gated) or the binding of small regulatory molecules (ligand-gated).

   Clinical importance of ligand-gated ion channels:
   Many anaesthetic drugs act on ligand-gated ion channels!

   - Nicotinic acetylcholine receptors (nAchR)
     - Pentameric ligand-gated ion channels opened by Ach and nicotine

- *Ach* is the endogenous neurotransmitter at these receptors
- *Nicotine* is a potentially toxic alkaloid derivative of tobacco. It mimics certain actions of Ach and was used to investigate the physiology of the ANS; at low doses being stimulatory at nAchRs.

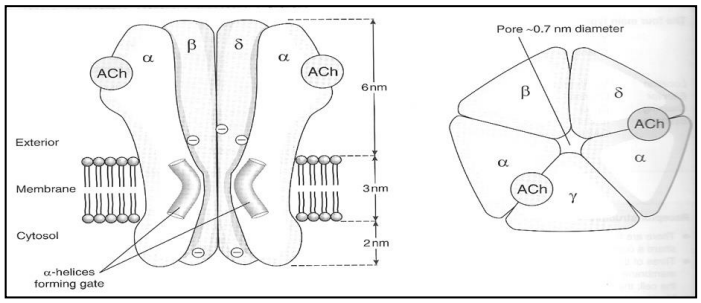


Figure 1: Schematic diagram of the nicotinic acetylcholine receptor

- Location of nAchR:
  - At the neuromuscular junction of skeletal muscle
  - On postganglionic neurons at the ganglion of the autonomic nervous system
- Composed of 5 subunits (2 $\boldsymbol{\alpha}$, and one β, δ and $\boldsymbol{\varepsilon}$ *or* $\boldsymbol{\gamma}$) which span the entire membrane and create a central pore
- Ach binding sites are located on the 2 $\boldsymbol{\alpha}$ subunits; both of which need to be bound to Ach to trigger conformational change which 'opens' the pore and allows for the influx of calcium and sodium ions, and the efflux of potassium
- The $\boldsymbol{\alpha}$ subunit is also the binding site of acetylcholine receptor agonists and antagonists

❖ $\boldsymbol{\gamma}$-Aminobutyric acid (GABA) type A receptors
- These receptors are a pentameric collection of 3 different subunits: $\boldsymbol{\alpha}$, β, $\boldsymbol{\gamma}$
- $GABA_A$ receptors mediate fast postsynaptic inhibition
- These ion channels are selectively permeable to chloride. The resultant influx of chloride ions results in hyperpolarization which inhibits action potential firing, and hence decreased excitability
- Apart from binding its endogenous ligand GABA, which is the main inhibitory neurotransmitter in the CNS, the receptor has multiple binding sites which facilitate the action of GABA. Drugs acting here include benzodiazepines (which augment the opening induced by GABA), barbiturates, propofol, etomidate, and possibly inhaled anaesthetic agents. The different anaesthetic agents are thought to interact with different subunits of the receptor.

❖ N-Methyl-D-aspartate (NMDA) receptors
- Named after its potent exogenous agonist
- Ligands: excitatory neurotransmitter glutamate which requires glycine as a co-agonist. Glutamine and glycine bind different subunits of the receptor
- Activation leads to opening of a cation-selective ion channel, resulting in the influx of Na+ and Ca++ and efflux of K+ ions. Also, activation results in an increased response to glutamate by a positive feedback mechanism, thought to lead to a hyper-excitable state ('wind-up') whereby repeated stimuli cause increasing degrees of pain sensation and expansion of sensory neurons involved in pain pathways
  - NMDA receptors can be phosphorylated by the serine/threonine kinases namely protein kinase C (PKC),

PKA and calcium/calmodulin-dependent protein kinase II, as well as tyrosine kinases
- In general, phosphorylation *enhances* NMDA receptor function

- Exhibits a voltage-dependent block by magnesium ions
- Antagonist: Ketamine

❖ 5-HT_3 receptors
- 5-hydroxytryptamine or *serotonin* is widespread throughout the body
- The 5-HT_3 receptor is the only of the 7 families of serotonin receptors which is an ion-gated channel (the rest are all G-protein coupled receptors)
- Activation results in excitatory effects mediated via ion channels which control the influx of Na^+ and K^+ ions
- Receptors are located in the CNS and PNS. Those located in the area postrema/vomiting centre in the medulla play an important role in nausea and vomiting
  - MOA of antagonists such as ondansetron and granisetron

Clinical significance of voltage-gated ion channels:

❖ Voltage-gated channels include Na+ channels and Cl^- channels
- Local anaesthetics work by blocking voltage-gated Na+ channels

b. Transporters

Transporters include exchangers, co-transporters, and ATP-driven ion pumps. Functionally, transporters are excitable membrane proteins that, after relatively selective binding of transporter ions, undergo conformational changes to allow physical movement of ions across the membrane. Transporters facilitate active transport of certain ions against their electrochemical gradients, thus establishing and maintaining transmembrane electrochemical gradients.

Primary active transporters are dependent on the hydrolysis of ATP to ADP by ATP-ase. ATPase-coupled ion pumps are typified by the Na^+-K^+-ATPase which actively transports three Na^+ extracellularly against its concentration gradient in exchange for two K^+.

Na^+-K^+-ATPase is a heterodimer made up of an α subunit and a β subunit. There are 3 α subunits ($\alpha_{1,2,3}$) and 3 β subunits described. α_1 is found in most cellular membranes; α_2 in muscle, heart, adipose tissue and brain; and α_3 in heart and brain. β_1 is found in most tissues, but *absent* astrocytes, vestibular cells and glycolytic fast-twitch muscles which only contain β_2 subunits.
The β subunit is a glycoprotein with a single membrane-spanning domain and 3 extracellular glycosylation sites. The α subunit is where the ion transport occurs; it is thought to span the cell membrane 10 times and when Na^+ binds to the α subunit, ATP also binds and is converted to ADP. The phosphate generated binds to the phosphorylation site on the α subunit, and this causes a conformational change in the protein with Na^+ extrusion into the ECF. K^+ then binds extracellularly which dephosphorylates the α subunit, and as it returns to its previous conformation, K^+ is released into the cytoplasm.

Clinical importance of Na^+-K^+-ATPase

❖ Drugs acting on these ion pumps alter the resting membrane potential
- Digitalis
  - Inhibits Na^+-K^+-ATPase which has particular importance in the heart where the Na^+-K^+ exchange is replaced by Na^+-Ca^{++}

exchange. This results in increased intracellular calcium which increases myocardial contractility

Secondary active transporters couple movement of one ion against its electrochemical gradient to the movement of another down its electrochemical gradient.
Examples of secondary active transporters are the Na^+-K^+-$2Cl^-$ co-transporter and Na^+/H^+ exchanger.

2. Metabotropic or G-protein coupled receptors (GPCRs)

GPCRs are the largest family of membrane receptors and mediate most cellular responses to hormones, ions, photons, neurotransmitters and other stimuli. Approximately 40-50% of drugs on the market target GPCRs!

Simplistically speaking, these membrane-bound proteins have a serpentine structure and traverse the cell membrane 7 times with alternating intracellular and extracellular loop regions. The binding of a ligand to the extra-cellular side results in activation of a G-protein on the cytosolic side, which in turn activates intermediate messengers, to bring about an often amplified intra-cellular change. Signal amplification occurs as a result of the intra-cellular messenger potentially being reused after the initial stimulus. The extracellular surfaces and ligand-binding sites have significant structural divergence, and a wide variety of ligand-binding "modes" will have quite different effects, from full or partial agonism through to antagonism.

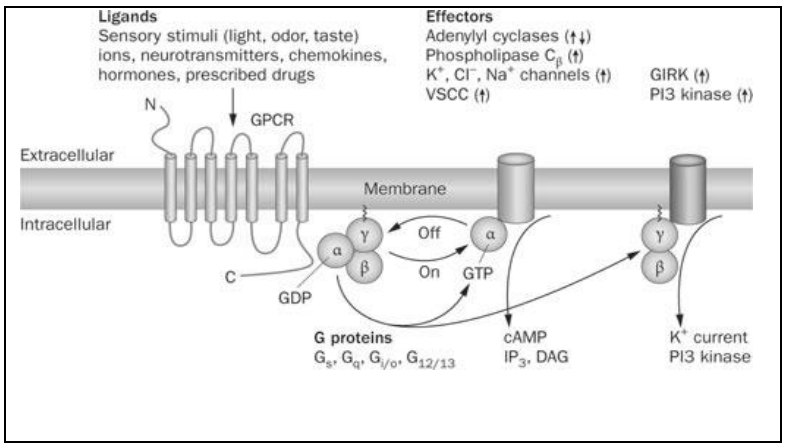


Figure 2: schematic illustration of the effects of GPCRs

The larger metabotropic G-protein coupled receptors are heterotrimeric having three subunits: $\boldsymbol{\alpha}$, β & $\boldsymbol{\gamma}$. In the inactive form, the $\boldsymbol{\alpha}$-subunit is bound to GDP. When a ligand binds and activates the receptor, GDP is exchanged for GTP. The $\boldsymbol{\alpha}$-GTP subunit dissociates from the β$\boldsymbol{\gamma}$ dimer and either activates or inhibits an effector protein (most commonly adenyl cyclase or phospholipase C). Thereafter the $\boldsymbol{\alpha}$-subunit acts as GTPase and regenerates the inactive $\boldsymbol{\alpha}$-GDP subunit. $\boldsymbol{\alpha}$-GDP and the β$\boldsymbol{\gamma}$ dimer recombine to form the heterotrimeric receptor complex and the process may start over. The β$\boldsymbol{\gamma}$ dimer also has the potential to exert an effect in certain scenarios such as activating G-protein-regulated inwardly rectifying potassium (GIRK) channels, and plays a role in anchoring regulatory kinases to the cell membrane. Apart from activating potassium channels, intracellular G-protein subunits may also mediate calcium, sodium and chloride channel activity in certain tissues.
Now, just as you're thinking this all makes perfect sense, please remain aware of the fact that these G-proteins relay signals from over a thousand ligands, and the effects within the cell are widely varied. They are not a simple "two-state switch", and many ligands may even activate more than one class of GPCRs with differing effects.
Also to be considered is the concept of desensitization, which can be defined as "waning of physiologic responsiveness to a drug over time". Here stimulation of receptor pathways leads to phosphorylation of specific regions of the receptor (due to the activation of kinases such as protein kinase A, G-protein coupled receptor kinase or protein kinase C) which prevent further receptor and/or second messenger activity. Phosphorylation by a G-protein coupled receptor kinase may also result in coupling to arrestin. Arrestin is a signaling and regulatory protein that

promotes the activation of extracellular signal-regulated kinases (ERK), prevents the activation of G-proteins, and promotes the internalization of receptors through clathrin-coated pits. Acute desensitization is termed tachyphylaxis.

A host of different sub-classes of G-proteins exist, 4 of which (determined by the α-subunit) will be discussed here, namely G_s, G_i, G_K and G_q.

i. **<u>G_s: stimulate adenyl cyclase (AC)</u>** resulting in increases in cyclic adenosine 3',5'-monophosphate (cAMP). cAMP exerts the effects of protein synthesis, gene activation or alterations in permeability through the stimulation of the R unit of protein kinase A and the phosphorylation of proteins involved in muscle contraction.

cAMP is broken down by phosphodiesterases (this should turn on a little 'pharmacology light bulb' in your head regarding actions and effects of the phosphodiesterase inhibitors!).

<u>Clinical importance of G_s-proteins:</u>

- β-adrenergic receptors (β_1, β_2, β_3)
  - All act via G_s proteins to increase the cellular level of cAMP which is indirectly responsible for inotropic and electrophysiologic effects
  - β_1 receptors:
    - heart: increase rate and force of cardiac contraction – cAMP activates a specific PKA that phosphorylates several important cardiac ion channels including L-type Ca^{2+} channels, Na^+ channels, voltage-dependent K^+ channels, and Cl^- channels. Ultimately, the increased intracellular calcium enhances inotropy
    - adipose tissue: lipolysis
    - kidney (juxtaglomerular apparatus): renin release
  - β_2 receptors:
    - vascular smooth muscle (sm): relaxation occurs as cAMP-dependent protein kinase phosphorylates myosin light chain kinase
    - bronchial sm; intestinal sm; bladder sphincter: relaxation
    - salivary glands: watery secretion
    - liver: glycogenolysis
    - pancreas: increased insulin & glucagon secretion
    - *<u>To highlight the complexity of the GPCR system</u>: β_2 adrenergic receptors can activate G_i as well as G_s proteins, which differentially regulate AC*
  - β_3 receptors:
    - adipose tissue: lipolysis

- cAMP second messengers are downregulated by specific phosphodiesterase proteins. Phosphodiesterase III inhibitors indirectly increase cAMP levels by inhibiting its breakdown
  - Phosphodiesterase is also inhibited by methylxanthines such as caffeine and theophylline

- The cholera toxin alters the alpha subunit of the G_s protein which inhibits its GTPase activity, resulting in prolonged stimulation of adenylyl cyclase

- 5-HT_4 receptors:
  - located in the CNS, GIT, bladder heart
  - stimulation mainly results in increased GIT motility
  - Pharmacologic agonists include metoclopramide
- Other receptors include DA_1, H_1, H_2, HT_2, glucagon, ACTH, LH, FSH, VIP, GHRH, TRH and prostacyclin

ii. **<u>G_i: inhibits AC</u>** thus reducing intracellular cAMP. G_i also activates phospholipase A_2 which activates the arachidonic acid cascade

Clinical importance of G_i-proteins:

- α_2-adrenergic receptors
  - located in presynaptic adrenergic nerve terminals where they inhibit the release of NA and Ach (negative feedback loop). Also found on platelets, lipocytes and smooth muscle
  - receptors are further subdivided into:
    - α_{2A} : responsible for central regulation of BP, sympathetic activity, pain processing and alertness
    - α_{2B} : cause vasoconstriction
    - α_{2C} : thought to be involved in behavioural responses
  - Pharmacologic agonists include clonidine and dexmedetomidine

- Opioid receptors:
  - Opioid receptor nomenclature is based on their founding anatomical location and pharmacological profile, namely, *m*orphine (*m*u or MOP), *k*etocyclazocine (*k*appa or KOP), and vas *d*eferens (*d*elta or DOP). Each of these receptors is the product of a single gene and are regarded as the 'classical' opioid receptors because they are sensitive to the antagonist naloxone. There is a fourth receptor, the nociception (NOP) receptor whose endogenous ligand is nociception/orphanin FQ (N/OFQ), which is currently classified as a non-opioid member of the family because it is *not* sensitive to naloxone.
  - Most opioid drugs used clinically work through the MOP receptor
  - All four opioid receptor subtypes bind to G_i proteins to reduce the activity of adenylyl cyclase, reducing levels of intracellular cAMP. This also results in closure of voltage-sensitive calcium channels (VSCCs) and stimulation of potassium efflux which effectively hyperpolarizes the cell. The overall effect is reduced neurotransmission and reduced neuronal excitability, as well as inhibition of neurotransmitter release.
  - I would highly recommend a really great article on this topic in BJA Education from 2014 titled *Opioid Receptors* (see references)

- Muscarinic acetylcholine receptors (mAchR) M_2 & M_4
  - Agonists for all subtypes include Ach and muscarine
  - Antagonists vary between subtypes but include atropine, scopolamine, diphenhydramine, ipratropium, chlorpromazine, haloperidol and mamba toxin
  - M_2 receptors:
    - In the CNS, M_2 receptors play an important role in muscarinic agonist-mediated tremor and temperature control
    - In the PNS, receptors in the heart control cardiac myocyte contractility resulting in slowing of the heart rate, reduced contractile force of the atrium, and reduced conduction velocity of the AV node

      - **A subtype of G_i receptors, the G_k class or atrial muscarinic M_2 receptors, is linked to K^+ channels**
      - These receptors were initially thought to occur only in the atria, however it is now known that they also exist in the ventricles but in lower concentrations
      - The more important signaling mechanism than changes in cAMP at the cardiac atrial location is opening of an inwardly rectifying K^+ channel in the plasma membrane. Here the $\beta\gamma$ subunit activates the K^+ channel
      - Cardiac adenosine receptors are also coupled to this channel

- M_4 receptors:
  - located in the CNS where overall effects are inhibitory

❖ $GABA_B$ receptors
- Reduced formation of cAMP after receptor activation ultimately results in inhibition of voltage-gated calcium channels. There is thus reduced neurotransmitter release
  - Pharmacological agonists include baclofen

❖ 5-HT_1 receptors
- act mainly as inhibitory pre-synaptic receptors and result in neural inhibition and vasoconstriction
- subtypes 5-HT_{1A} & 5-HT_{1D}
  - The 5-HT_{1D} receptor agonist, sumatriptan, is used in the treatment of migraine

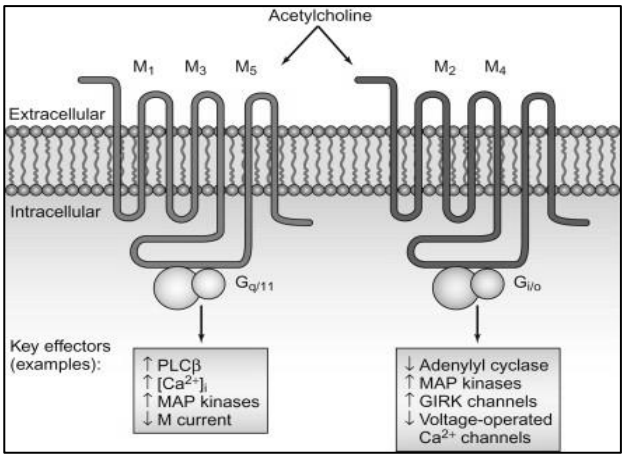
❖ The pertussis toxin interferes with the alpha subunit of G_i

iii. **<u>G_q: activate phospholipase C (PLC)</u>** which controls the breakdown of phosphoinositides to form inositol 1,4,5-triphosphate (IP_3) and diacylglycerol (DAG). IP_3 causes calcium release from the endoplasmic reticulum which then causes membrane hyperpolarization or enzyme release. Calcium then binds calmodulin which further activates kinases. DAG causes activation of protein kinase C which has various cellular effects based on interactions with ion channels, transporters, glycolytic enzymes and transcription factors.

<u>Clinical importance of G_q-proteins:</u>

❖ α_1- adrenergic receptors:
- vascular & bladder sm sphincter: contraction
- iris: contraction
- intestinal sm: relaxation (but sphincter contraction)
- uterus: variable
- salivary glands: viscous secretion
- liver: glycogenolysis
- pancreas: decreases secretion of enzymes, insulin & glucagon

❖ mAchR M_1, M_3 & M_5
- Agonists for all subtypes include Ach and muscarine
- Antagonists vary between subtypes but include atropine, scopolamine, diphenhydramine, ipratropium, chlorpromazine, haloperidol and mamba toxin
- M_1
  - Receptors are located in the autonomic ganglia, salivary glands and stomach
  - They predominate in the CNS, being found in the hippocampus, cerebral cortex, and striatum. The functioning of these receptors are critical for memory processes.
- M_3
  - widespread location and play an important role in the contraction of smooth muscle (airway, ileum, iris, bladder)
  - saliva and insulin secretion
  - Induce emesis
  - Paradoxical vasodilation is due to increased nitric oxide production by vascular endothelial cells
- M_5
  - Located in vascular endothelium (especially cerebral vessels), and CNS



- 5-HT_2 receptors
  - 5-HT_{2A} mediates excitatory effects on platelets and sm

3. <u>Catalytic receptors</u>

<u>Tyrosine kinases</u> are located within the cell membrane. Insulin, epidermal growth factor (EGF) and platelet-derived growth factor (PDGF) act through this receptor, as well as many drugs. These receptors have a single membrane-spanning domain with an intracellular tyrosine kinase (TK) domain. Binding of a ligand results in autophosphorylation or cross-phosphorylation of the TK domain. The result is the production of transcription factors in the nucleus which alter gene expression. <u>Serine/threonine kinase receptors</u> have a similar mechanism of action.

<u>Clinical importance of tyrosine kinases:</u>

- Insulin receptor:
  - Found in primary target tissues: liver, muscle, adipose tissue
  - Receptors bind insulin with high specificity and affinity
  - The insulin receptor consists of 2$\boldsymbol{\alpha}$ and 2β subunits which are covalently linked. The $\boldsymbol{\alpha}$ subunit is extracellular and is essentially the 'binding' site. The β subunit spans the membrane and contains a tyrosine kinase. Binding of insulin to the $\boldsymbol{\alpha}$ subunits activates the receptor through a conformational change which bring the β subunits closer together, facilitating mutual phosphorylation of tyrosine residues and activation of tyrosine kinase. Activated tyrosine kinases phosphorylate docking proteins, insulin receptor substrate 1 to 6 (IRS-1 to IRS-6), which results in a complex cascade of further phosphorylation within the cell which represent insulin's second message and results in:
    - translocation of glucose transporters (esp. GLUT 4) to the cell membrane which results in increased glucose uptake
    - increased glycogen synthase activity and increased glycogen formation
    - effects of protein synthesis, lipolysis, lipogenesis
    - activation of transcription factors that enhance cell growth and division
  - Glucocorticoids lower the affinity of insulin receptors for insulin
  - Growth hormone increases the affinity

<u>Guanylyl cyclases:</u> Just as adenylyl cyclase catalyses the formation of cAMP, guanylyl cyclases are a family of enzymes that catalyze the formation of cGMP which then activates cGMP-dependent kinases with a number of physiologic effects. Guanylyl cyclases exist in 2 forms as illustrated by these two examples:

- ❖ Receptors for atrial natriuretic peptide (ANP) have an extracellular amino terminal domain (the receptor portion) and a single transmembrane domain. The cytoplasmic portion has guanylyl cyclase catalytic activity which increases intracellular cGMP.
- ❖ The other form of guanylyl cyclase is intracellular and is activated by nitric oxide (NO) and NO-containing compounds.

4. Intracellular receptors

When certain ligands such as steroid hormones, thyroid hormones or retinoic acid bind receptors within the cell (either within the cytoplasm or the cell nucleus), a receptor-hormone complex is formed. This complex interacts with DNA via zinc fingers and binds to certain genes. This increases the transcription of encoded mRNAs which are in turn translated in the ribosomes. The result is increased production of proteins that alter cell function.

The receptors are mostly similar in structure, having cysteine-rich DNA-binding domain; a ligand-binding domain near the carboxyl terminal; and a nondescript amino terminal region.

Conclusion:

I hope this talk highlights the relevance of this highly complex system with specific reference to anaesthesia and the drugs we use on a daily basis.
Good luck!

References:

1. Ganong W.F. Review of Medical Physiology. 22nd Edition, 2005
2. Miller R.D. Miller's Anaesthesia Seventh Edition, 2010
3. Dubyak GR. Ion homeostasis, channels, and transporters: an update on cellular mechanisms. *Adv Physiol Educ* 28: 143-154, 2004
4. Zito K, Scheuss V. NMDA Receptor Function and Physiological Modulation. 2009
5. Rosenbaum DM, Rasmussen SGF. The structure and function of G-protein-coupled receptors. *Nature* Vol 459: 356-363, 2009
6. McDonald J, Lambert, DM. Opioid receptors. *Continuing education in anaesthesia, critical care and pain,* 2014
7. Peck TE, Hill SA, Williams M. Pharmacology for Anaesthesia and Intensive Care. 2nd Edition, 2006
8. Kaplan, JA. Kaplan's Cardiac Anaesthesia. 7th Edition, 2017: 217-220
9. Arcache M. Receptors and second messenger systems. Part I Anaesthetic Refresher Course, 2009
10. Radford, H. Signal transduction – receptors and second messenger systems. Part I Anaesthetic Refresher Course, 2011
11. Milner A, Welch E. Applied pharmacology in Anaesthesiology and Critical Care. 2012

The Pharmacokinetics of Intravenous Opioids

Dr Sean Chetty

Deputy Head
Department of Anaesthesiology and Critical Care
University of Stellenbosch & Tygerberg Hospital

Opioid agents are fundamental to the practice of modern anaesthesia. Opioids work by interacting with specific opiate receptors. Of these, the mu receptors appear most important in terms of analgesia and the respiratory depressant opioid effects. The mu receptor type consists of the least two subtypes, μ_1 which mediates analgesia and μ_2 which mediates respiratory depression, bradycardia, and physical dependence. The G protein second messenger system is activated by opioid agonists binding to the receptor. Subsequently, activated G protein subunits alter membrane permeability (increasing K^+ in decreasing Ca^{2+} conductance) which has the effect of hyperpolarizing the membrane thus inhibiting neuronal activity.

Sites of opioid effects include: medulla, spinal cord, spinal trigeminal nucleus, and periaquaductal grey area, which is an integration modulation site from peripheral nerves to the central neuraxis.

Opioids are weak bases (pKa 6.5-8.7). In solution, they dissociate into ionised and unionised fractions, the relative proportions depend upon the pH of the solvent and their pKa. The unionised fraction is more diffusible than ionised form. High lipid solubility facilitates opioid transport into the biophase or site of action. Consequently, high lipid solubility confers a more rapid onset of action. Drugs with high lipid solubility, high unionised fraction or low protein binding in the plasma, demonstrate large volumes of distribution. Most opioids are extensively distributed in the body and their volumes of distribution exceed total body water. Small intravenous doses of short acting opioids (like alfentanil, sufentanil or fentanyl) produce short durations of action because plasma (and brain) concentrations remain above the threshold for therapeutic action for only a brief period, as the drug rapidly redistributes from the CNS to other tissues. Larger doses produce longer durations of action because plasma concentrations remain above the threshold at the completion of drug redistribution and depend upon the slower elimination process to be reduced below the threshold level.
Opioids are metabolised mainly in the liver to both active and inactive compounds that are excreted in urine and bile. Morphine and other opioids are excreted partly in the bile as water-soluble glucuronides. In the gut, these glucuronides are metabolised by the normal gut flora to the parent opioid compound and reabsorbed (entero-hepatic recirculation). Highly lipid soluble opioids, for example fentanyl, may diffuse from the circulation into the stomach mucosa and lumen, where they are ionised and concentrated because of the low pH. Later, gastric emptying and reabsorption from the small intestine may produce a secondary peak effect (gastro-enteric recirculation). Extra-hepatic metabolism is important for some opioids eg. the kidneys play a vital role in conjugating morphine, whereas blood and tissue esterases are responsible for remifentanil metabolism.

Opioids differ substantially in their durations of action. Explanations for these differences are complex and not always evident from their clearance and terminal half-lives. For example, an analgesic dose of morphine lasts longer than a dose of fentanyl producing an equivalent degree of analgesia; yet the half-life for morphine is shorter than fentanyl. In the case of morphine, its relatively long duration of action is a reflection of its relatively low lipid solubility and slow diffusion out of CNS tissue. Once it enters blood it is effectively cleared from plasma.

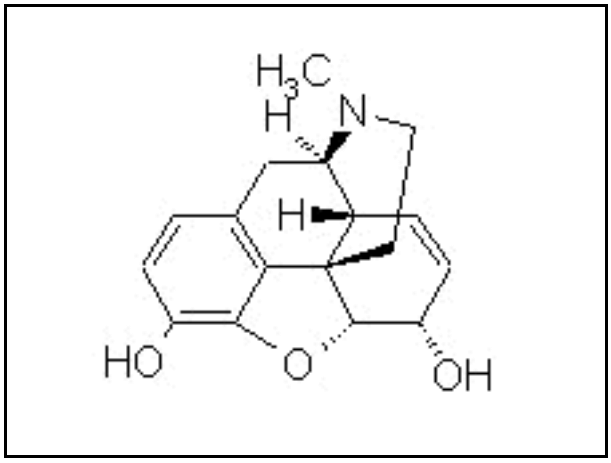
Much of the currently available pharmacokinetic data for the opioids are derived from single-dose studies completed in healthy volunteers. Single-dose studies however fail to predict the pharmacokinetic parameters that are seen after long-term infusions of these agents because of the multi-compartment behavior of the parent drug and its metabolites. Results of studies completed in healthy volunteers can also not be extrapolated to patients who are critically ill (as in the case of ICU patients and in many patients who will present for emergency surgery) because the alterations in the volume status, plasma protein binding, and end-organ function that occur in this population will affect drug bioavailability, volume of distribution, and clearance of the drug from the body.

This review will aim to summarise the pharmacokinetic properties of the common opioid drugs currently used in Anaesthesia in South Africa.

Pharmacokinetics refers to the influence that the body has on a drug and is traditionally divided into four different areas: Absorption, Distribution, Metabolism and Elimination. When discussing the pharmacokinetics of intravenous opioids, absorption does not influence this discussion as intravenous administration of a drug negates the need for absorption of the drug.

INDIVIDUAL OPIOIDS

Morphine



Morphine is a naturally occurring phenanthrene derivative. It is the standard drug against which all other opioids are compared. Morphine can be given orally, intramuscularly (IM), intravenously (IV), subcutaneously (SC), rectally, epidurally and intrathecally. Intravenous administration should be titrated to effect (usually 1-2mg boluses), but the total dose remains similar to the intramuscular dose of 0.1-0.2mg/kg. Morphine may also be given epidurally at 10% and intrathecally at 1% of the parenteral dose.

Following bolus administration, onset time is relatively slow (15-30 minutes) because:

- morphine exhibits relatively low lipid solubility (about 2.5% of fentanyl)
- at physiological pH, morphine, a weak base with the pKa of about 8.0, is primarily ionized. The ionized form does not favor passage through the lipid membrane. It is only the 10%-20% of the molecules the are un-ionized, that are able to pass through the lipid membrane.

The volume of distribution of morphine is 3-5 L/kg.

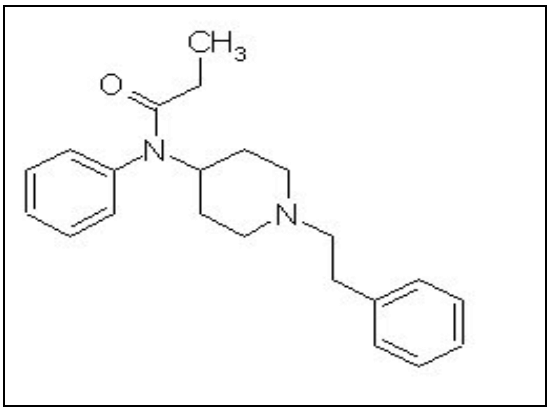
Morphine is extensively metabolised by the liver to morphine-3-glucuronide (M3G) (70%), morphine-6-glucuronide (M6G) (10%) and to sulphate conjugates. M6G is 10-20 times more potent than morphine and is normally excreted in urine. It accumulates in renal failure and accounts for increased sensitivity to morphine.

Neonates are more sensitive than adults to morphine due to their reduced hepatic conjugating capacity. In the elderly, owing to reduced volume of distribution, the peak plasma level of morphine is higher compared to younger patient.

The main effects of morphine are mediated through the mu receptors. It is a potent analgesic with good sedative and anxiolytic properties. It may cause euphoria, dysphoria and hallucination. It produces respiratory depression and cough suppression. It has minimal effect on the cardiovascular system but may cause vasodilation, bradycardia and hypotension. Nausea and vomiting are common side effects. Histamine release may lead to a rash, itching and bronchospasm (in susceptible patients). Meiosis is common. Tolerance and dependence may develop.

Elimination halftimes for morphine following a bolus administration is about 1.7-4.5 hours. The drug has a relatively high plasma clearance rate of 15-40ml/kg/minute. This imples that there is also an extrahepatic clearance mechanism involved in its elimination. The most likely mechanism is renal.

Fentanyl



Fentanyl is a synthetic phenylpyperidine derivative. It is 100 times more potent than morphine. When given in small doses (1-2 microgram/kg), it has a rapid onset and a short duration of action (30 min). Such doses are used intravenously for pain associated with shorter/minor surgery. In small doses it has little sedative effect. Higher doses are used to obtund the sympathetic response to laryngoscopy and intubation. Fentanyl has also been used to augment the effects of local anaesthetics in spinal and epidural analgesia at 10-25 microgram and 25-100 microgram doses respectively.

This drug is 500 times more lipid soluble than morphine, consequently it is rapidly and extensively distributed in the body (volume of distribution 4 L/kg).

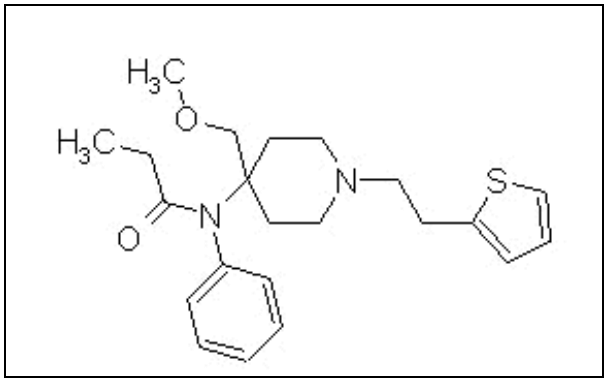
Fentanyl is also a weak base at physiological pH meaning that only about 10% of molecules are un-ionized.

Following parenteral administration, plasma and CNS concentrations fall below an effective level during the rapid distribution phase (for small doses [1-2 microgram/kg]). This is explained by rapid redistribution from brain to other compartments such as skeletal muscle and fat.

However, following prolonged administration or with high doses, its duration of action is significantly prolonged. In these circumstances, the distribution phase is complete while the plasma concentration is still high. Recovery from the effect of the drug then depends on its slow elimination from the body (terminal half life 3.5 hours). Fentanyl is predominantly metabolised in the liver to norfentanyl, which is inactive. The clearance of the drug is 10 – 20 ml/kg/min. The metabolite, norfentanyl, is excreted in the urine over few days.

Many properties of fentanyl are similar to morphine. It produces respiratory depression in dose-dependent manner. Large doses (50- 100 microgram/kg) have been used for cardiac surgery to obtund the metabolic stress response. At such high doses, sedation is profound and unconsciousness tends to occur. Muscular rigidity of the chest wall may also affect ventilation at these high doses.

Sufentanil



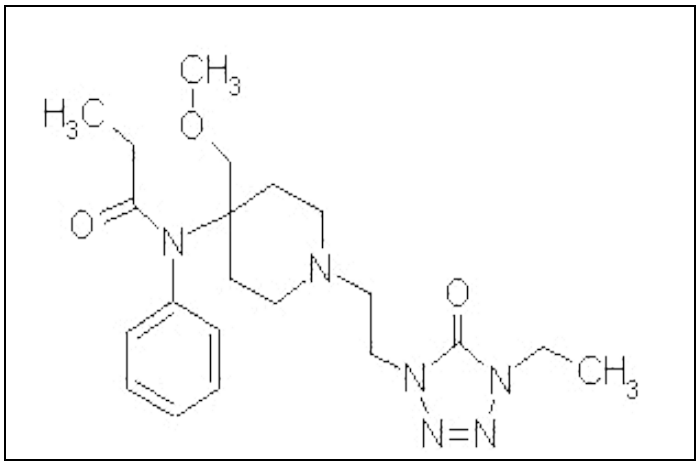
Sufentanil is a synthetic opioid that is very similar to fentanyl. Sufentanil is a weak base with a pKa of about 8, indicating that at physiological pH, about 20 percent of the sufentanil molecules are un-ionized, and are therefore able to readily traverse membrane lipid bilayers. The pharmacokinetics are best represented by a multicompartment (3 compartment) model.

IV bolus sufentanil has a rapid onset of action, similar to fentanyl, with rapid redistribution resulting in a relatively short-term duration of action. Elimination halftime is about 2.7 hours and and is associated with high clearance (13ml/kg/minute) which is primarily dependent on hepatic blood flow. The duration of action is shorter than for fentanyl although sufentanil is about 5-10 times more potent.

The increased sufentanil potency compared to fentanyl is probably related the greater lipid solubility and and higher receptor affinity of the molecule. Higher receptor affinity means that to achieve a certain percentage occupancy of the receptor a lower drug concentration is required.

"Context sensitive" halftimes (which are those following termination after continuous infusion) are shorter for sufentanil compared to alfentanil or fentanyl. The decline in plasma drug concentration will depend on both redistribution to peripheral compartments and metabolism.

Alfentanil



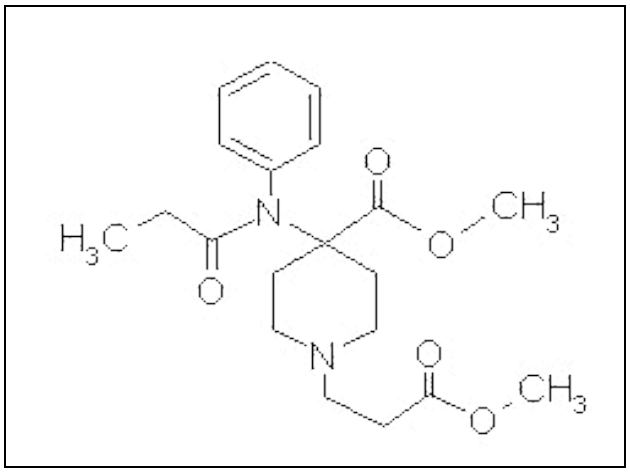
Alfentanil is a synthetic phenylpiperidine derivative structurally related to fentanyl. It has 10-20% of its potency. Alfentanil may be administered intravenously as either a bolus or by continuous infusion. Bolus doses (10microgram/kg) are useful for short-term analgesia and attenuation of the cardiovascular response to intubation. Continuous infusions (0.5-2.0 micrograms/ kg/min) are used in the intensive care unit for sedation in patients on mechanical ventilation or for TCI anaesthesia.

Although it has much lower lipid solubility than fentanyl, more alfentanyl is present as an unionised form compared to fentanyl (89% compared to 9%) after intravenous administration; consequently, its onset of action is more rapid. Also, because of its lower lipid solubility, less alfentanil is distributed to muscles and fat. Hence, its volume of distribution is relatively small and more of the dose remains in blood from which it can be cleared by the liver. Even though alfentanil has a lower clearance rate, this is more than offset by its reduced volume of distribution and its half-life is relatively short.

Most effects of alfentanil are similar to fentanyl but with quicker onset and shorter duration of action.

The drug clearance is 4-9 ml/kg/minute. Alfentanil metabolites are inactive although numerous. There are significant individual differences in clearance rates and this may reflect different gene expression levels of all the cytochrome P450 isoforms responsible for alfentanil metabolism, namely CYP3A4 (cytochrome P450 3A4 enzyme activity).

Remifentanil



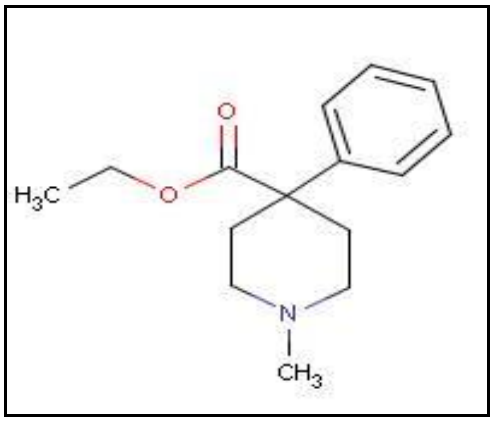
Remifentanil is a synthetic phenylpiperidine derivative of fentanyl with similar potency but is ultra short-acting. It is unique among the various opioids used in anaesthesia because of its susceptibility to ester hydrolysis. There are two ester-linkages in the molecule, one of which is rapidly cleaved by nonspecific plasma and tissue enzymes (different from pseudocholinesterase which hydrolyses drugs such as succinylcholine and mivacurium). Remifentanil is less soluble than other opioids. It has a very rapid onset of action, similar to Alfentanil.

A range of infusion rates (0.05- 2.0 microgram/kg/min) can be used during the maintenance of anaesthesia with controlled ventilation.

As stated above, Remifentanil is rapidly broken down by non-specific plasma and tissue esterases and this results in a short elimination half life of 3-10 minutes. This rapid ester hydrolysis limits its duration of action so significantly that, following its redistribution from the brain, its plasma half-life is extremely short. In fact, the plasma half-life is so short that there is very little redistribution to other compartments. It is context insensitive, in that the half-life, clearance and distribution are independent of duration and strength of infusion.

The special degradative pathway that it undergoes, is responsible for elimination halftimes on the order of 10-25 minutes, which is much less than in other intravenous agents.

Pethidine



Pethidine is a synthetic phenylpyperidine derivative and was originally developed as an antimuscarinic agent. Pethidine can be administered SC (50-100 mg), IM (50-100 mg) or IV (25-100 mg). The doses can be repeated every 4 hours.

It is 30 times more lipid soluble than morphine. The drug is metabolised in the liver by ester hydrolysis to norpethidine and pethidinic acid, which are excreted in the urine, and therefore accumulate in renal

failure. At higher concentrations, norpethidine can produce hallucination and convulsions. Pethidinic acid is an inactive compound. Pethidine is often used for labour analgesia. It readily crosses placenta, and a significant amount passes to the foetus over several hours.

There are some pharmacological differences from morphine. It produces a tachycardia, dry mouth and less marked meiosis. However, a significant fall in blood pressure may occur when pethidine is administered to the elderly or to hypovolaemic patients. It may however produce less biliary tract spasm than morphine.

Pethidine is absolutely contraindicated in patients on monoamine oxidase inhibitors (MAOI), as serious side effects like hypotension or hypertension, hyperpyrexia, convulsions or coma may occur. The underlying mechanism for these side effects is not clear but may involve the reduced metabolism of pethidine by MAOI and pethidine's effect on turnover of 5- hydroxytryptamine in the brain.

Table 1: Pharmacokinetics of commonly used opioids

	Morphine	Fentanyl	Sufentanil	Alfentanil	Remifentanil	Pethidine
pKa	8.0	8.4	8.0	6.5	7.1	8.5
Unionised at pH 7.4 (%)	23	9	20	90	68	5
Plasma protein bound (%)	30	84	93	90	70	40
Terminal half life (hrs)	3	3.5	2.7	1.6	0.06	4
Clearance (ml/min/kg)	15-30	0.8-1.0	13	4-9	30-40	8-18
Volume of distribution (L/kg)	3-5	3-5	2.5-3.0	0.4-1.0	0.2-0.3	3-5
Octanol/H_2O partition coefficient	1.4	813	1778	145	17.9	39

References

1. Barr J, Donner A. Optimal intravenous dosing strategies for sedatives and analgesics in the intensive care unit. Crit Care Clin 1995;11(4):827–47.
2. Benet LZ, Kroetz DL, Sheiner LB. The Dynamics of Drug Absorption, Distribution and Elimination. In, Goodman and Gillman's The Pharmacological Basis of Therapeutics,(Hardman, J.G, Limbird, L.E, Molinoff, P.B., Ruddon, R.W, and Gilman, A.G.,eds) TheMcGraw-Hill Companies, Inc.,1996, pp. 3-27
3. Calvey TN, Williams NE eds. Analgesic Drugs. In: Principles and Practice of Pharmacology for Anaesthetists, 4th Edn. Blackwell Science, 2001;196-228
4. Correia, M.A., Drug Biotransformation. in Basic and Clinical Pharmacology, (Katzung, B. G., ed) Appleton-Lange, 1998, pp 50-61.
5. Devlin JW, Roberts RJ. Pharmacology of Commonly Used Analgesics and Sedatives in the ICU: Benzodiazepines, Propofol, and Opioids. Anesthesiology Clin 29 (2011); 567–585
6. Dolin, S. J. "Drugs and pharmacology" in Total Intravenous Anesthesia, pp. 13-35 (Nicholas L. Padfield, ed), Butterworth Heinemann, Oxford, 2000
7. Hall LG, Oyen LJ, Murray MJ. Analgesic agents. Pharmacology and application in critical care. Crit Care Clin 2001;17(4):899–923, viii.
8. Hutton P, Cooper GM eds. Opioid Analgesics In: Fundamental Principles and Practice of Anaesthesia. Martin Dunitz Ltd, 2002;621-626
9. Katzung, B. G. Basic Principles-Introduction, in Basic and Clinical Pharmacology, (Katzung, B. G., ed) Appleton-Lange, 1998, pp 1-33
10. McDonald J, Lambert D: Opioid receptors. Continuing Education in Anaesthesia, Critical care & Pain 2005; Vol5 Number 1: 22-25.
11. Stoelting, R. K. and Miller, R.D. Intravenous Anesthetics, in Basics of Anesthesia, 4th edition, pp. 58-69, Churchill-Livingstone, 2000.
12. Stoelting, R.K., "Pharmacokinetics and Pharmacodynamics of Injected and Inhaled Drugs", in Pharmacology and Physiology in Anesthetic Practice, Lippincott-Raven Publishers, 1999, 1-17.
13. Trescot AM, Datta S, Lee M, et al. Opioid pharmacology. Pain Physician 2008; 11(2 Suppl):S133–53.

Urea and Creatinine

Dr Michelle Casey

UCT Dept of Anaesthesia & Perioperative Medicine

Urea and Creatinine Plasma levels

In South Africa, urea is measured in mmol/L, and serum creatinine is measured in μmol/L, whereas in the United States and most European countries both urea and creatinine are measured in mg/dL.

The normal range of urea in adults is 2.1 - 7.1 mmol/L. The normal creatinine range in adults is 64 - 104 μmol/L in males and 49-90 μmol/L in females (National Health Laboratory Services). Further wide inter-individual variability does exist as per the tables below.

Increased Urea levels	**Decreased Urea levels**
High Protein diet	Low protein diet
Dehydration	Malnutrition/starvation
GI bleed	Parenchymal liver disease
Catabolic processes (eg. Exercise, fever, burns)	Congenital urea cycle enzyme deficiencies
Drugs (corticosteroids, certain tetracyclines)	Pregnancy
Decreased renal excretion	
Increased hepatic urea synthesis	
Rhabdomyolysis	
Hypoperfusion (eg. congestive cardiac failure, shock)	
Decreased blood volume (eg. haemorrhage)	
Urinary tract obstruction	

Increased Creatinine levels	**Decreased Creatinine levels**
Ingestion of cooked meat	Increased age
Drugs (eg. Creatine supplements, trimethoprim, cimetidine)	Decreased dietary intake (eg. Vegetarian diet)
Dietary intake - Level may be slightly elevated in the evening due to meat consumption	Women (less muscle mass)
Rhabdomyolysis	Reduced muscle mass (eg.Amputees, muscle wasting)
Sickle cell disease, nephrotic syndrome (increased tubular creatinine secretion)	Pregnancy
High muscle mass (eg. bodybuilders)	

Urea Production

Urea (60Da) is a primary metabolite derived from dietary protein and tissue protein turnover, and facilitates the excretion of excess nitrogen in the form of ammonia from the body. A diet containing 70g protein, will produce on average, 12g of urea in 24 hours.

Dietary protein, as well as endogenous protein derived from physiological catabolism and degradation of blood for example, is converted to peptides and amino acids in the small intestine. More than 90% of these amino acids are absorbed into the blood stream and transported to the liver where they undergo further deamination and transamination. The unabsorbed peptides, as well as recycled urea, are converted into ammonia by the gut flora present in the colon. Ammonia then diffuses via the portal circulation, and reaches the liver where it is converted to nitrogenous biomolecules (eg. nucleotides). The excess ammonia enters the urea cycle.

Urea synthesis occurs predominantly in the periportal hepatocytes, and to a lesser degree in the enterocytes. In the mitochondria, NH^{4+} combines with CO_2 to form carbamoyl phosphate. This step is facilitated by the enzyme, carbamoyl phosphate synthetase, and forms the rate-limiting step. In essence, the urea cycle consists of 4 main steps (Myles, 2003). The first step of the urea cycle occurs in the matrix of the mitochondria, where ornithine transcarbamylase transfers a carbamoyl group from carbamoyl phosphate to ornithine to form citrulline. The subsequent steps occur in the cytosol. The second step is catalyzed by argininosuccinate synthetase, which converts citrulline and aspartate to argininosuccinate. In the third step, arginosuccinate is converted to fumarate and arginine via argininosuccinate lyase. The last step is catalyzed by arginase, where arginine is converted into urea and ornithine is regenerated. 3 ATP are consumed during the entire cycle. Urea is then excreted in the urine, thereby facilitating the excretion of excess nitrogen.

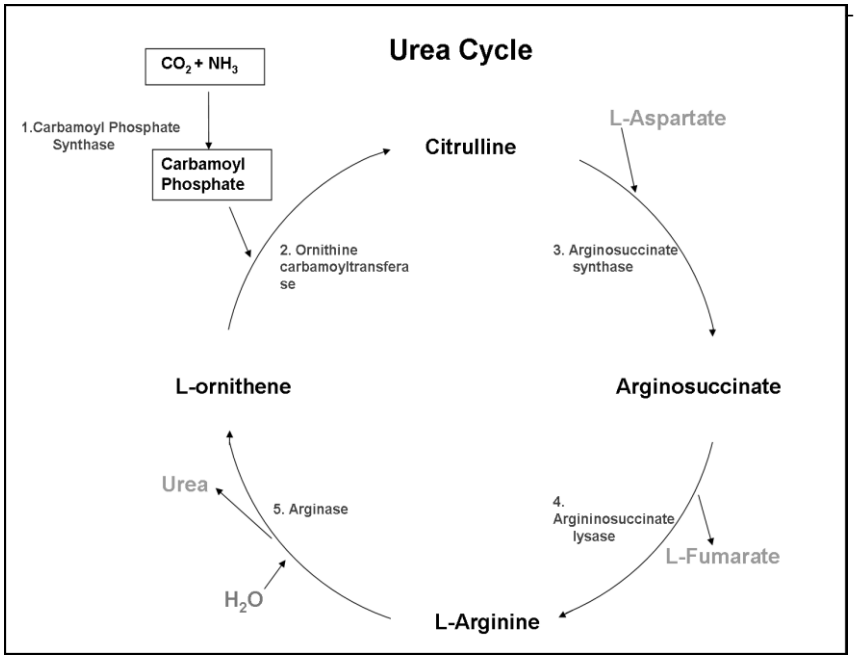


Image obtained from https://upload.wikimedia/org/wikipedia/en/6/64/Urea_cycle_1.png

Creatinine production

Creatinine (113Da) is primarily a byproduct of skeletal muscle creatine metabolism. Creatinine production is therefore a reflection of an individual's lean body mass and is produced at a constant rate. It is also derived from dietary protein (to a lesser extent), as heat from cooking converts the creatine in meat to creatinine.

Biosynthesis of skeletal muscle creatine starts with the production of Guanidinoacetic acid (GAA) from transfer of the amidino group of arginine to glycine. This initial reaction occurs primarily in the kidneys, but can also occur to a lesser extent in the mucosa of the small intestine and in the pancreas. GAA is

then transported to the liver, where it is methylated by S-adenosyl methionine to form creatine. The creatine is then released into the circulation, where 90% is taken up and stored by muscle tissue. The stored creatine is phosphorylated to phosphocreatine, an energy-storage molecule, in a reaction catalyzed by creatine phosphokinase.

Muscular phosphocreatine and creatine is broken down spontaneously/non-enzymatically to creatinine at a constant rate. Overall, 1.1% of the body's creatine stores and 2.6% of the phosphocreatine stores are converted to creatinine in a day (Wyss, 2000). The newly formed creatinine then diffuses out of the muscle into the blood stream, where it is excreted by the kidney.

Creatinine clearance

Creatinine clearance is used to estimate the glomerular filtration rate (GFR) (Inker LA, 2015), which serves as an index of kidney function. Inulin remains the ideal marker to calculate GFR as it is freely filtered through the glomeruli and is not secreted, metabolized or reabsorbed by the kidney. It is however not routinely available, is expensive and is difficult to handle. Other more ideal markers include iohexal, iothalamate, DTPA or EDTA.

Tubular secretion of creatinine by the organic cation secretory pathway in the proximal tubule accounts for 10-20% of urinary creatinine in patients with a normal GFR, and higher values in those patients with a lower GFR, due to hypersecretion from the functional nephron tubules. This results in overestimation of the GFR by 10-20% in normal individuals and increasingly higher levels of overestimation as renal disease worsens.

Traditionally, the use of older colorimetric techniques to measure serum creatinine resulted in an overestimation of the s-creatinine by 10-20% due to the contribution of non-creatinine chromogens (eg. acetone, pyruvate and ascorbic acid) to the final serum creatinine concentration. The error secondary to tubular creatinine secretion was thus balanced by the laboratory error of overestimation of the s-creatinine. This measurement error has however largely been abolished due to national standardization of serum creatinine assays to creatinine reference materials, resulting in the creatinine clearance once again overestimating the patient's true GFR by 10-20%. Other errors that can occur when calculating creatinine clearance are:

1). Inaccuracies in urine collection. Incomplete collection leads to an underestimate of creatinine excretion. A 24hour urine collection is used to determine the urine creatinine concentration, as a shorter period for collection gives less accurate results, as creatinine excretion varies throughout the day. GFR is calculated by dividing 24hour urine creatinine by s-creatinine.

2). Increases in creatinine secretion in the proximal tubule which occurs when there is a decrease in the creatinine filtration and excretion, resulting in an elevated s-creatinine level. Increased creatinine secretion results in an overestimate of the true GFR, as the creatinine secretion is far higher than the creatinine filtration.

3). Increased extrarenal creatinine degradation. In patients with advanced renal failure (s-creatinine >530 μmol/L) there is an overgrowth of intestinal bacteria and bacterial creatininase activity. This in turn leads to an increase in extrarenal creatinine breakdown. This causes the s-creatinine level to be lower than that expected from the GFR, and falsely elevates the creatinine clearance.

4). Creatinine clearance fails to detect mild renal disease, as there is compensatory hypertrophy and hyperfiltration in the residual normal nephrons, resulting in a minimal decrease in the GFR.

Calculation of creatinine clearance

If the contribution of creatinine secretion by the proximal tubule is ignored, then all the filtered creatinine (GFR x serum Creatinine concentration (SCr)) will be excreted in the urine (Urine Creatinine concentration (UCr) x Urine flow rate or volume (V)).

i.e. GFR x SCr = UCr x V
$\therefore$ GFR = [UCr x V]/SCr

Eg. Calculate the creatinine clearance in a 50kg female, body surface area of 1.5 kg/m^2.
SCr= 90 μmol/L
UCr= 7500 μmol/L
Urine volume= 1.1 L/day

GFR= [UCr x V]/SCr
= [7500 x 1.1L/day]/90= 91.7 L/day

To convert L/day to ml/min: convert L to ml by multiplying by 1000, then divide by the number of minutes in a day (1440) i.e. 91.7 X 1000/1440= 64 ml/min.

These results should be normalized to a body surface area of 1.73 m^2. For example, if the above patient has a BSA of 1.5 kg/m^2:

Creatinine clearance X 1.73/BSA= [64 ml/min X 1.73]/1.5
∴adjusted creatinine clearance = 74ml/min per 1.73m^2

The normal range of values for creatinine clearance is:
Female: 95 ± 20 ml/min/1.73 m^2
Male: 120 ± 25 ml/min/1.73m^2

In patients with a stable s- creatinine level, the GFR can be estimated from empirically derived estimation equations using the serum creatinine value. These include the Modification of Diet in Renal Disease (MDRD) equation and the 2009 Chronic Kidney Disease Epidemiology (CKD-EPI) equation. This avoids the need to collect 24hour urine specimens. Another alternative in patients with moderate to advanced renal disease (s-creatinine >220umol/L), is to take the average of the creatinine and urea clearances: GFR= Creatinine clearance + Urea clearance/2.

Urea Excretion

A small amount of urea (<0.5g/day) is lost through the gastrointestinal tract, lungs and skin. A larger amount of urea may be excreted in sweat during exercise. The bulk of urea (10g/day) however, is excreted by the kidney.

Urea transport in the kidney plays a vital role in the urine concentrating mechanism, by creating an osmotic gradient in the medullary pyramids that facilitates the production of a concentrated urine in the collecting ducts (Barrett, 2010). Urea transport (UT) proteins mediate urea transport, most likely by facilitated diffusion. Four isoforms of the urea transport protein (UT-A1 to UT-A4) exist in the kidney. UT-B is found in the descending vasa recta and the red blood cells.

Urea is filtered across the glomerulus to the proximal tubule. The amount of urea entering the proximal tubule is controlled by the GFR. The amount of urea filtered also depends on the dietary protein intake, with a high protein diet increasing the ability of the kidney to concentrate the urine. The urea concentration increases to a concentration 50% greater than that of plasma in the first 75% of the proximal convoluted tubule (PCT). This is due to reabsorption of water out of the PCT secondary to salt transport.

Two types of loops of Henle occur in the kidney. The long looped occur in the juxtamedullary nephrons, whilst the short looped (absence of the thin ascending limb) occur in the cortical nephrons. All portions of the short loops are permeable to urea, however the amount and direction of urea movement depends on the individual's diuretic status.

The intraluminal concentration of urea rises in the thin descending limbs, due to water loss resulting in an increase in the urea: water ratio. Water reabsorption is driven by the hypertonic medullary interstitium which results from the movement of urea out of the terminal inner medullary collecting duct (IMCD). Urea permeability across the IMCD increases when there is an increase in osmolality, or by the presence of antidiuretic hormone (vasopressin). Both hypertonicity and ADH increase the urea permeability of the IMCD by regulating the urea transporters, UT-A1 and UT-A3, thereby increasing the osmotic gradient further and increasing water reabsorption to result in a more concentrated urine (Klein JD, 2011). Urea remains in the interstitial space and is not removed by the circulation, as the

vasa recta act as countercurrent exchangers. Urea is also secreted into the lumen of the thin limbs under antidiuretic conditions.

The urea concentration increases further in the thin ascending limb of the loop of Henle, due to the gradient for urea secretion provided by urea reabsorption from the IMCD. The gradient decreases as the thin ascending limb ascends and the driving force to move urea into the tubular lumen decreases. The urea concentration reaches an equi-osmolar level with the surrounding interstitium by the beginning of the medullary thick ascending limb. Despite the thick ascending limb having a lower urea permeability, there is still a mild increase in the intraluminal urea concentration from the beginning of the thick ascending limb to the distal convoluted tubule.

The distal convoluted tubule (DCT) has a relatively low urea permeability, however some of the urea is reabsorbed, thereby decreasing the urea concentration from 110% of the filtered load to 70% by the start of the proximal cortical collecting duct. Both the cortical and medullary collecting ducts have a low urea permeability. As discussed earlier, the urea permeability of the inner medullary collecting duct can be increased under hypertonic conditions, and under the influence of ADH. 30-50% of the filtered urea will ultimately be excreted in the urine (Weiner ID, 2015)

Creatinine excretion

Creatinine is excreted unchanged, primarily by the kidneys. As creatinine is a small molecule (113 Da), it is freely filtered across the glomerulus. Unlike urea, it is not reabsorbed. As discussed earlier under 'creatinine clearance', secretion of creatinine by the proximal tubules accounts for 10 - 20% of the total amount of creatinine excreted in the urine. Extrarenal creatinine excretion can also occur in patients with advanced kidney disease (eGFR<15ml/min/1.73cm^2), as intestinal bacterial overgrowth with increased creatininase activity leads to extrarenal creatinine excretion and a falsely low serum creatinine level. In normal steady state conditions, the rate of creatinine excretion = creatinine production in the steady state.

- **Estimated creatinine excretion (mg/24hrs**) = 1115.89 + (11.97 X weight in kg) – (5.83 x age) - (60.18 x phosphorous in mg/dL) + (52.82 if black) – (368.75 if female)
- **Adult male < 50 years:** daily creatinine excretion is 20-25mg/kg (177-221 μmol/kg) of lean body mass
- **Adult female <50 years**: 15-20mg/kg (133-177 μmol/kg) of lean body mass
- **50-90 years:** 10mg/kg in men (50% decrease in creatinine excretion, which correlates to a decreased muscle mass).

Urea and creatinine levels in pregnancy

During pregnancy, the glomerular filtration rate (GFR) rises secondary to an increase in the cardiac output and renal blood flow (80% increase above non-pregnant levels), leading to an increased glomerular plasma flow. This increase in GFR occurs within the first month of conception, and peaks at 40-50% above baseline by the early 2nd trimester. The level then decreases towards term, but remains elevated above non-pregnant levels up to 1 week postpartum. All values have returned to baseline by 6-8 weeks postpartum (Odutayo & Hladunewich, 2012).

The increase in GFR results in a decrease in the serum creatinine and urea levels in pregnant women (Cheung & Lafayette, 2013). Blood urea levels fall to 2.9 – 3.9 mmol/L. (Thadhani RI, 2016). The s-creatinine level decreases by an average of 35 μmol/L, to a normal range of 35-70 μmol/L. Vigilance is therefore required, as renal injury may present with only small fluctuations in s-creatinine levels, despite a significant decline in the patient's renal function.

A study conducted in patients with preeclampsia revealed that although preeclampsia was associated with a decrease in GFR by up to 40%, only a minority of patients had a serum creatinine level that exceeded normal non-pregnant ranges (Hladunewich MA, 2008). The serum creatinine value only exceeded normal non-pregnant ranges in those patients whose GFR was decreased by greater than 50% below the pregnant control mean of 149 ml/min/1.73 m^2.

Urea and creatinine levels in the elderly

The amount of creatinine production decreases with age, primarily due to a decrease in muscle mass. The decrease in lean body mass subsequently results in a decrease in the glomerular filtration rate. A longitudinal study conducted in 254 men over a duration of 23 years, demonstrated a decline in GFR of 7.5 ml/min for each decade of life (Lindeman RD, 1985).

Creatinine clearance continues to decrease substantially with age to 43ml/min/1.73m^2 (>75 years) versus 144ml/min/1.73m^2 (18-40 years). The physiological decline in renal function associated with the normal aging process, as well as a higher incidence of disease that affects the kidney, also contribute to a decline in renal function. Physiological changes that occur with normal aging include both micro-anatomical changes (decrease in nephron number with secondary compensatory hypertrophy of the remaining functional nephrons, nephrosclerosis) and macro-anatomic changes (renal cyst development, cortical atrophy). (Rule Ad, 2016). There are currently no proven interventions or therapy available to prevent the age-related decline in GFR.

Serum creatinine values do not always reflect the age-related decline in GFR. The serum creatinine values remain stable with healthy aging, as the decrease in GFR is cancelled out by the decrease in the individual's lean body mass,

Interestingly, urea excretion increases with aging, with an increase in fractional urea excretion of 40% (>75 years) vs 24% (18-40years) in volume contraction, and 65% (>75 years) vs 53% (18-40 years) in volume expansion. The reason for the increased fractional excretion of urea has been suggested to be due to the lower reabsorption of urea at the distal tubules. The increase in urinary urea excretion, as well as a low protein diet typically associated with the elderly population, explain the normal serum urea values found in the elderly, despite a decrease in their GFR (Musso, Alvarez Gregori, Jauregui, & Macias Nunez, 2012).

References

1. Barrett, K.K, Barman, S.M., Boitano, S., Brooks, H.L. (2010). *Ganong's review of medical physiology* (Vol. 23): McGrwa-Hill.
2. Cheung, K. L., & Lafayette, R. A. (2013). Renal physiology of pregnancy. *Adv Chronic Kidney Dis, 20*(3), 209-214. doi:10.1053/j.ackd.2013.01.012
3. Hladunewich MA, Myers BD, Derby GC, et al. (2008). Course of preeclamptic glomerular injury after delivery. *Am J Physiol Renal Physiol, 294*, F614.
4. Inker LA, Perrone RD. (2015). Calculation of creatinine clearance. www.uptodate.com
5. Klein JD, Blount MA, Sands JM. (2011). Urea transport in the kidney. *Compr Physiol, 1*, 699-729.
6. Lindeman RD, Tobin J, Shock NW. (1985). Longitudinal studies on the rate of decline in renal function with age. *J Am Geriatr Soc, 33*, 278.
7. Musso, C. G., Alvarez Gregori, J., Jauregui, J. R., & Macias Nunez, J. F. (2012). Creatinine, urea, uric acid, water and electrolytes renal handling in the healthy oldest old. *World J Nephrol, 1*(5), 123-126. doi:10.5527/wjn.v1.i5.123
8. Myles, B. (Producer). (2003). The Urea Cycle. Retrieved from www.tamu.edu/faculty/bmiles/lectures/urea.pdf
9. Odutayo, A., & Hladunewich, M. (2012). Obstetric nephrology: renal hemodynamic and metabolic physiology in normal pregnancy. *Clin J Am Soc Nephrol, 7*(12), 2073-2080. doi:10.2215/CJN.00470112
10. Rule Ad, Glassock RJ. (2016). The aging kidney. www.uptodate.com
11. Thadhani RI, Maynard SE. (2016). Renal and urinary tract physiology in normal pregnancy. www.uptodate.com
12. Weiner ID, Mitch WE, Sands JM. (2015). Urea and Ammonia Metabolism and the Control of Renal Nitrogen Excretion. *Clin J Am Soc Nephrol, 10*(8), 1444-1458. doi:10.2215/CJN.10311013
13. Wyss, M., Kaddurah-Daouk, R. (2000). Creatine and creatinine metaboslim. *Physiological reviews, 80*(3), 1107-1213.

The Regulation of Chloride

Dr Kotie Bester

Division of Paediatric Anaesthesia
UCT Dept of Anaesthesia & Perioperative Medicine

Chloride is the ionized form of the chlorine atom. In solution, chloride ions are bound by the protic ends of water molecules. Chloride is the major anion in the extracellular space. It contributes to osmotic pressure and osmotic gradients between extra- and intracellular spaces; it contributes to electrical potentials and electroneutrality; and it helps maintain hydration of cells, tissues and secretions. It has a major role to play in acid-base balance. Table 1 lists the normal values for chloride concentration in specified fluids.

Fluid	Concentration (mmol/l)
Plasma	95-107
Cerebrospinal fluid	116-122
Urine	110-250

Table 1: *Normal values for chloride concentration*

Most chloride is excreted via the kidney, with small additional amounts excreted via sweat and faeces. Excessive sweating may lead to significant losses via this route. Similarly vomiting and diarrhoea may cause significant losses. Table 2 lists potential causes of hypo- and hyperchloraemeia.

Hypochloraemia	Hyperchloraemia
Defective renal tubular absorption	Dehydration
Vomiting	Excessive intake
Diarrhea	Swallowing of sea water
Respiratory acidosis	Iatrogenic (hypertonic saline administration)
Loop diuretics	Aspirin intoxication
	Compensation for respiratory alkalosis
	Renal failure
	Congestive heart failure
	Cystic fibrosis

Table 2: *Causes of hypochloraemia and hyperchloraemia*

Renal regulation of chloride concentration

Glomerular filtration

As with most electrolytes, chloride is filtered extensively by the glomerulus, and then reabsorbed by the tubular system. Ninety-nine per cent of filtered chloride is reabsorbed. Most of this chloride follows sodium, with a few exceptions.

Reabsorption

Sodium movement causes an *electrochemical gradient* that draws chloride along passively – these chloride ions use a paracellular pathway to pass from the tubular lumen directly into the interstitial space.
Water reabsorption via osmotic forces additionally creates a concentration gradient for chloride reabsorption.
Active transport of chloride ions is a “secondary active” process. The most important of these processes is co-transport with sodium across the luminal border.

Reabsorption: Proximal tubule
Up to 65% of the filtered sodium, and slightly less chloride, is absorbed in the proximal lumen. *Sodium* provides a *co-transport* mechanism for chloride reabsorption. However, in the early proximal tubule sodium co-transport also aids reabsorption of glucose, amino acids, bicarbonate, etc., and some sodium is exchanged for hydrogen. In the latter half of the proximal lumen, a *concentration gradient* for chloride has developed because sodium (and water) reabsorption was coupled to many other solutes, not just chloride. Therefore, in the latter half of the proximal tubule, chloride also diffuses down its concentration gradient via the paracellular pathway directly into the interstitial space. (Table 3)

First half of proximal tubule		**Modifying drugs / hormones**
Sodium reabsorption	Exchange with hydrogen Co-transport with glucose, amino-acids, bicarbonate (Co-transport with chloride – see next)	Carbonic anhydrase inhibitors decrease hydrogen secretion (and sodium reabsorption). Angiotensin II stimulates the Na-K-ATPase.
Chloride reabsorption	*Co-transport* with sodium, driven by Na-K-ATPase pump	
Second half of proximal tubule		
Chloride reabsorption	*Diffusion* due to concentration gradient	
Loop of Henle		
Chloride reabsorption	*Na-2Cl-K co-transporter*, driven by Na-K-ATPase pump Chloride moves from cells into the interstitium via "ClC-Kb channels"	Loop diuretics inhibit Na-2-Cl-K co-transporter.
Early distal tubule		
Chloride reabsorption	*Na-Cl-co-transport*, driven by Na-K-ATPase pump	Thiazide diuretics inhibit Na-Cl co-transporter.
Late distal tubule and collecting duct		
Chloride reabsorption	*Diffuses* with sodium, driven by Na-K-ATPase pump	Aldosterone stimulates synthesis of ENaC's (and activation of the Na-K pump). Spironolactone inhibits aldosterone. Vasopressin increases water reabsorption.
Hydrogen secretion	Intercalated cells actively secrete H^+	
Chloride exchange	Bicarbonate exits the cell (towards the interstitium) in exchange for chloride via the anion exchanger "Band 3"	

Table 3: *Tubular handling of chloride*

Reabsorption: Loop of Henle
About 25% of the filtered chloride, sodium and potassium get absorbed in the loop of Henle – mostly in the thick ascending limb via the *Na-2Cl-K co-transporter*. This co-transporter uses the potential energy created by the Na-K-ATPase at the basolateral membrane of the luminal cells. The Na-K-ATPase pumps sodium out of the luminal cell into the interstitial space, creating a concentration gradient for sodium. This concentration gradient draws sodium into the cell from the tubular lumen. (This mechanism is also used in the proximal tubule.)

Reabsorption: Distal tubule
Chloride moves with sodium again, using *Na-Cl-co-transport*, driven by the concentration gradient created by the Na-K-ATPase pump. Chloride channels in the basolateral membrane provide an avenue for chloride diffusion into the interstitial space.

Reabsorption: Late distal tubule and collecting duct
The Na-K-ATPase pump creates a concentration gradient for sodium and chloride **diffusion**. Sodium movement is via the epithelial sodium channels (ENaC). Aldosterone regulates this action.

Some chloride secretion may take place when chloride passively follows secreted hydrogen into the tubular lumen. Chloride enters the tubular cell from the interstitial space in exchange for bicarbonate ions via an anion exchanger called Band 3.

Renal chloride channels

A number of chloride channels have been described that promote chloride reabsorption in different parts of the nephron. Examples are a channel mediating reabsorption in the loop of Henle, a channel in the basolateral membrane of the distal nephron, and a channel in vesicles of the proximal tubule. These channels not only mediate chloride reabsorption, but also have an important role in acidification of intracellular vesicles and cell volume regulation.

Extra-renal chloride channels

Chloride channels have many functions elsewhere.

Inner ear: Chloride channels in the inner ear help maintain a high K+ concentration that is essential for normal hearing.

Nervous tissue: The cell membranes of nerve cells have chloride channels (ClC). These channels permit negatively charged chloride ions to move between the cell and the extracellular space. Changing the concentration of chloride ions can decrease excitability of nerve cells.
GABA-A receptors are also chloride channels. (GABA-B receptors are G-protein coupled receptors, and therefore very different.) GABA-A receptors need a concentration gradient for chloride to drive an inward current when they open. ClC's may aid in creation of this concentration gradient. If down regulation of ClC's decrease the ability to create a concentration gradient for chloride, opening the GABA-A receptor may cause outward flux of the anions leading to depolarization (excitation). This may play a role in neuropathic pain, as well as during early development when excitation can be seen with GABA-A stimulation.
The glycine receptor is another chloride channel present in nerve cells: it has inhibitory actions in the brain stem and spinal cord.

Epithelial cells: Cystic fibrosis transmembrane conductance regulator (CFTR) is an ATP-gated ion channel for the conduction of chloride and thiocyanate ions down their concentration gradients across epithelial cell membranes. Fluid transport in lungs, the pancreas and other organs rely on chloride movement, and CFTR dysfunction causes thickened (insufficiently hydrated) secretions. Sweat glands need CFTR to transport sodium, chloride and thiocyanate during reabsorption of these ions. This is the basis for the "sweat test" in the diagnosis of cystic fibrosis.

Hormones affecting chloride regulation

Chloride concentration is a function of chloride mass and water volume. Water handling will therefore affect chloride concentration. It is clear that chloride follows sodium to a great extent, and sodium movement will also affect chloride movement. It follows that the following hormones will (often indirectly) affect chloride concentration.

Vasopressin (anti-diuretic hormone) increases water absorption. Its effect is seen in the collecting ducts, where it stimulates V_2 receptors. Water channels (aquaporin-2) are then inserted into the luminal surfaces of the principal cells.

Angiotensin II increases sodium reabsorption through actions on arterioles, but also by directly stimulating the basolateral Na-K-ATPase pumps throughout the tubular system, and stimulating sodium-hydrogen exchange in the proximal tubule.

Aldosterone is secreted in response to hyperkalaemia, severe hyponatraemia, or hypotension. Angiotensin II increases the release of aldosterone. Aldosterone increases sodium reabsorption in the distal convoluted tubule and collecting duct through increased expression of ENaC's and increased activity of the Na-K-ATPase pumps.

Atrial- and brain natriuretic peptides (ANP and BNP) have effects on glomerular filtration, but also inhibit sodium reabsorption in the renal tubules.

Electric forces and acid-base balance

Chloride anions play an important role in the neutralization of positive charges in the body. The concentration of anions (mainly chloride) is about 2% lower in plasma than in interstitial fluid. This is due to the *Donnan effect* (the phenomenon where equilibrium of charged particles across a semipermeable membrane is determined by both electrostatic and osmotic forces, i.e. charged proteins alter the movement of other ions).

Sodium and chloride are found in a specific proportional relationship in the body, and by ***Stewart's principles*** both act as strong ions. If either of the two ions is gained or lost in excess of the other, acid-base disturbance will result. If chloride is gained or lost in excess of sodium, metabolic acidosis or alkalosis will ensue respectively.

Changes in acid-base status due to chloride loss or gain can also be interpreted in terms of *renal compensation*. Ion concentrations in the extracellular fluid are generally adjusted via renal and respiratory mechanisms. Any chloride loss from the body will be offset by bicarbonate retention by the kidneys, and can therefore lead to an alkalosis.

Changes in acid-base status can affect ion concentrations: During primary respiratory alkalosis, the kidney will excrete bicarbonate and retain chloride in an attempt to compensate – this can lead to hyperchloraemia as part of metabolic compensation.

Chloride is a major determinant of the *anion gap.* The anion gap is often used to distinguish between different causes of acidosis.

$$\text{Anion gap} = Na^+ - Cl^- - HCO_3^-$$

A metabolic acidosis secondary to additional acids, e.g. lactic acidosis, increases the anion gap. Renal failure and renal tubular acidosis, where bicarbonate is lost and chloride is retained, causes an acidosis where hyperchloraemia is associated with a normal anion gap.

Interpretation of chloride levels: If free water is lost, sodium and chloride levels should rise proportionate to each other. The expected chloride level associated with a measured sodium level can be calculated as follows:

$$\text{Expected } Cl^- = (\text{normal } Na^+ / \text{measured } Na^+) \times \text{measured } Cl^-$$

(where normal Na+ is the midpoint of its reference interval)

Chloride in red blood cells

Other interactions relevant to chloride include the chloride shifts seen in red blood cells.

Hamburger effect: Chloride enters red blood cells (RBC) in exchange for bicarbonate ions via the anion exchange protein, Band 3. This results in a lower chloride concentration in venous blood than in arterial blood, and maintains a concentration gradient for CO_2 into the cell as follows:
CO_2 enters the RBC and forms H_2CO_3 (catalysed by carbonic anhydrase). This dissociates into H^+ and HCO_3^-. The hydrogen remains in the RBC, but the bicarbonate leaves the cell in exchange for chloride.

Haemoglobin and oxygen affinity: The Haldane effect, where oxygenation of haemoglobin displaces carbon dioxide increasing the removal of carbon dioxide, is an example of a change in haemoglobin's affinity for carbon dioxide. Chloride may act as an allosteric factor to regulate the affinity of haemoglobin for oxygen, and the chloride shift may play a role in this.

The Standard ECG
Why it looks the way it does

Dr Adriaan Myburgh

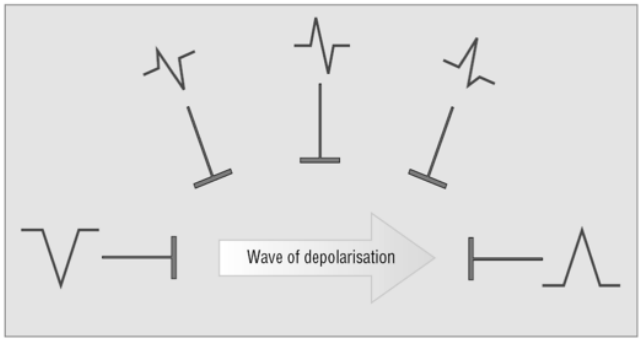
UCT Dept of Anaesthesia & Perioperative Medicine

1. Introduction:
The electrocardiogram (ECG) is a representation of the electrical events of the cardiac cycle.
Each event has a distinctive waveform.
The study of waveform can lead to greater insight into a patient's cardiac pathophysiology.

2. With ECGs we can identify:
Arrhythmias
Myocardial ischemia and infarction
Pericarditis
Chamber hypertrophy
Electrolyte disturbances (i.e. hyperkalemia, hypokalemia)
Drug toxicity (i.e. digoxin and drugs which prolong the QT interval)
Contraction of any muscle is associated with electrical changes called depolarization.
These changes can be detected by electrodes attached to the surface of the body.

3. Pacemakers of the heart:
SA Node - Dominant pacemaker with an intrinsic rate of 60 - 100 beats/minute.
AV Node - Back-up pacemaker with an intrinsic rate of 40 - 60 beats/minute.
Ventricular cells - Back-up pacemaker with an intrinsic rate of 20 - 45 bpm.

4. The ECG convention:
Standard calibration:
25 mm/s and a height representing 0.1 mV/mm
(Horizontally: One small box = 0.04 s, one large box = 0.20 s.
Vertically one large box = 0.5 mV)
Electrical impulse that travels towards the electrode produces an upright (positive) deflection



Wave of depolarisation. Shape of QRS complex in any lead depends on orientation of that lead to vector of depolarisation

Figure 1: Depolarisation wave

Impulse conduction: From the Sinoatrial node to AV node then down the bundle of His and bundle branches to reach the epicardium via the Purkinje fibers. This is represented as the P wave, the PR interval and the QRS complexes. The QRS represents ventricular depolarization and the T wave ventricular repolarization.
The PR interval represents atrial depolarization and delay in AV junction (AV node/bundle of His). This delay allows time for the atria to contract before the ventricles contract.

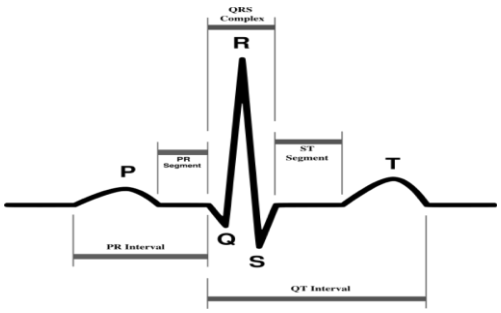


Figure 2: Normal ECG

The ECG leads measure the difference in electrical potential between two points, (bipolar leads) while the unipolar leads use one point on the body and a virtual reference point with zero electrical potential, located in the center of the heart.
The standard ECG has: 3 standard limb leads, 3 augmented limb leads and 6 pre-cordial leads.

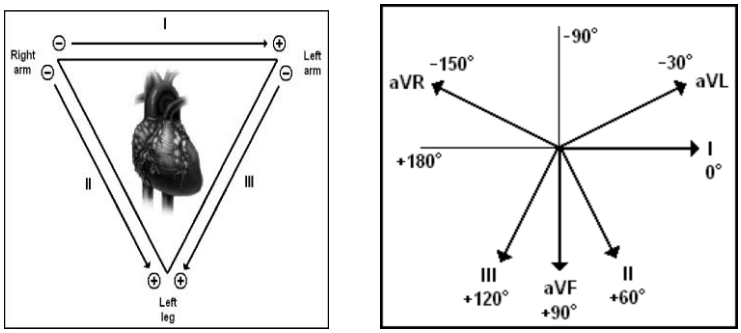
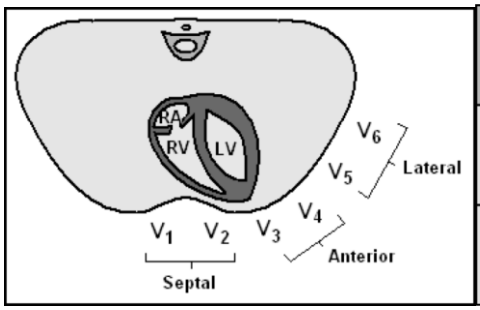


Figure 3: Conventional arrangement of limb leads

<u>The pre-cordial leads:</u>
The pre-cordial leads are arranged so that lead V4 lies in the midclavicular line 5th intercostal space. V6 lies lateral in the anterior axillary line. V5 lies between V4 and V6. V3 lies slightly to the right and above V4, as does V2, V1 lies to the right of V2 to the right of the sternum



I Lateral	aVR None	V_1 Septal	V_4 Anterior
II Inferior	aVL Lateral	V_2 Septal	V_5 Lateral
III Inferior	aVF Inferior	V_3 Anterior	V_6 Lateral

Figure 4: Arrangement of the pre-cordial leads.

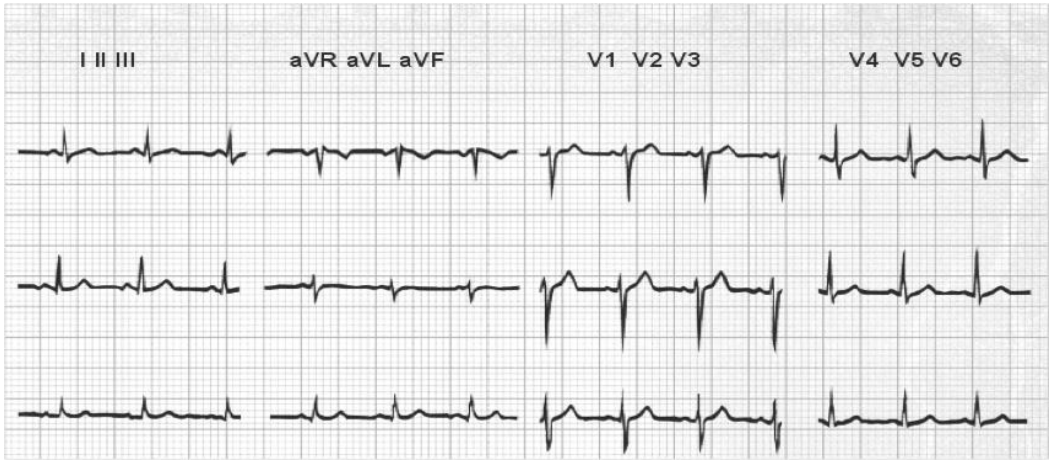


Figure 5: The normal ECG

Rule 1: PR interval should be 120 to 200 milliseconds or 3 to 5 little squares

Rule 2: The width of the QRS complex should not exceed 110 ms, less than 3 little squares

Rule 3: The QRS complex should be dominantly upright in leads I and II

Rule 4: QRS and T waves tend to have the same general direction in the limb leads

Rule 5: All waves are negative in lead aVR

Rule 6: The R wave must grow from V1 to at least V4.The S wave must grow from V1 to at least V3 and disappear in V6.

Rule 7: The ST segment should start iso-electric, except in V1 and V2 where it may be elevated

Rule 8: The P waves should be upright in I, II, and V2 to V6

Rule 9: There should be no Q wave or only a small Q less than 0.04 seconds in width in I, II, V2 to V6

Rule 10: The T wave must be upright in I, II, V2 to V6

4.2 The P wave:

- Always positive in lead I and II
- Always negative in lead aVR
- < 3 small squares in duration
- < 2.5 small squares in amplitude
- Commonly biphasic in lead V1
- Best seen in leads II

4.3 The QRS complex:

- Non pathological Q waves may be present in I, III, aVL, V5, and V6
- R wave in lead V6 is smaller than V5
- Depth of the S wave, should not exceed 30 mm
- Pathological Q wave > 2mm deep and > 1mm wide or > 25% amplitude of the subsequent R wave

4.4 Tall R in V1:

RV Hypertrophy
Posterior infarct
WPW syndrome
Right Bundle branch block
Normal in children and young adults

4.5 LV Hypertrophy:

S in V1+ R in V5 or V6 > 35 mm
A R wave of 11 to 13 mm (1.1 to 1.3 mV) or more in lead aVL is another sign of LVH

4.6 ST segment:

• ST segment is flat (isoelectric)
• Elevation or depression of ST segment by 1 mm or more is pathological
"J" (Junction) point is the point between QRS and ST segment

4.7 T Wave:

• Normal T wave is asymmetrical, first half having a more gradual slope than the second half
• Should be at least 1/8 but less than 2/3 of the amplitude of the R wave
• T wave amplitude rarely exceeds 10 mm
• Abnormal T waves are symmetrical, tall, peaked, biphasic or inverted.
T wave follows the direction of the QRS deflection.

4.8 QT Interval:

• Total duration of depolarization and repolarization
• QT interval decreases when heart rate increases
  - For HR = 70 bpm, QT<0.40 s.
  - QT interval should be 0.35- 0.45 s
  - Should not be more than half of the interval between adjacent R waves (RR interval).

5. Heart Rate

Rule of 300: Count the number of "big boxes" between two QRS complexes, and divide this into 300. (Smaller boxes with 1500) for regular rhythms.
Ten second rule: ECGs record 10 seconds of rhythm per page, count the number of beats present on the ECG . Multiply by 6 . For irregular rhythms.

6. The QRS Axis:

The QRS axis represents overall direction of the heart's electrical activity.
Abnormalities are due to:
Ventricular enlargement
Conduction blocks (i.e. hemiblocks)

Normal QRS axis from -30° to +90°.
-30° to -90° is referred to as a left axis deviation (LAD).
+90° to +180° is referred to as a right axis deviation (RAD).
Use QRS complex in leads I and aVF as they are perpendicular to one-another.
Determine if they are predominantly positive or negative.
The combination should place the axis into one of 4 quadrants.

The Pharmacology of Sedation

Dr Karmen Kemp

Division of Paediatric Anaesthesia
UCT Dept of Anaesthesia & Perioperative Medicine

Sedation is a continuum of an altered state of consciousness, ranging from minimal sedation to deep sedation. The objectives of sedation, as set out in the SASA guidelines, are to reduce anxiety, minimize physical discomfort and pain, minimize psychological trauma and maintain patient safety.

Adverse events may arise from poor patient selection, management and monitoring, but also from an inadequate knowledge of the pharmacokinetic and pharmacodynamics profile of the drugs given. Pharmacogenetic variability must also be considered.

Several classes of sedative drugs exist and can be used alone or in combination. When multiple drugs from different categories are used simultaneously, it must be considered that some drugs potentiate each other, such as benzodiazepines and opioids.

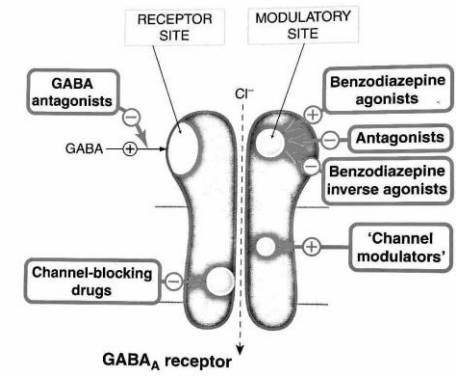
Midazolam and the α2-adrenoceptor agonists, Dexmedetomidine and Clonidine, are the most popular agents for procedural and intensive care sedation.

Benzodiazepines

Benzodiazepines consists of a 7 membered ring fused to an aromatic ring. It has 4 main substituent groups that confer some degree of selectivity and function. Different benzodiazepines are selected more for their pharmacokinetic profile than for their specific action.

Benzodiazepines are used as anticonvulsants, anxiolytics, sedatives, amnesiacs and provide some degree of muscle relaxation. It has no analgesic properties.
It binds selectively and reversibly to a regulatory site on the $GABA_A$ receptor. It acts allosterically to enhance the affinity of GABA to its binding site without influencing the binding of glutamate or glycine. It does not alter channel opening.

The $GABA_A$ receptor



The $GABA_A$ receptor is a ligand gated, pentamer with α,β and γ subunits. These subunits exist in a multitude of variations and determine function and drug selectivity. This ligand gated chloride channel is mainly located in the central nervous system where it mediates fast inhibitory synaptic transmission. Peripheral binding sites do exist but their function is unknown.

Action of benzodiazepines

α1 subunits in the $GABA_A$ receptor is more associated with sedative, amnesic and anticonvulsant properties and the α2 subunit with anxiolytic and central muscle relaxant properties.

- **Cardiovascular system**

Benzodiazepines cause virtually no cardiovascular depression. Care should still be taken in the elderly and patients with low cardiovascular reserve especially if used in combination with other sedatives or cardiac depressant drugs.

- **Respiratory system**

Benzodiazepines cause slight respiratory depression that is exaggerated when combined with other sedatives especially opioids. Benzodiazepines and opioids display synergism with regards to sedation and respiratory depression. It also causes decreased muscle tone that can lead to airway obstruction in the obtunded patient.

- **Sleep physiology**

Benzodiazepines reduce rapid eye movement (REM) sleep as well as the slow wave phase of sleep. Long term reduction of REM sleep causes irritability and anxiety and a rebound increase in REM sleep is seen after the cessation of benzodiazepines.

Midazolam

The most commonly used Benzodiazepine due to its fast onset (20-40 minutes oral; 1-3 minutes intravenous) and short duration of action. Midazolam is metabolised by cytochrome P450 enzymes through glucuronide conjugation. Metabolites have very little activity and are excreted renally. Advance age, renal function and liver function affect the pharmacokinetic factors of midazolam.

Midazolam is water soluble at a pH<4, but become lipid soluble in vitro, pH>7,4 due to closure of the imidazole ring. The elimination half life is 1.5-2 hours.

Midazolam causes paradoxical excitement in about 1% of adults and up to 10% of younger children (see Pharmacogenomics).

Tolerance and dependence develop with the prolonged use of Midazolam with a withdrawal syndrome characterized by agitation, anxiety, tremors and dizziness.

Flumazenil is a selective benzodiazepine receptor antagonist that can be used as an antidote to Benzodiazepine overdose. It has a duration of action of 2 hours and repeat doses might have to be given for longer acting benzodiazepine overdose.

Remimazolam

A new drug to be watched is Remimazolam. a new innovation, using the parent compound Midazolam and incorporating a carboxylic ester linkage to make it suitable for metabolism by non specific tissue esterases in the blood. The onset of action is predictably the same as intravenous Midazolam – 1-3 minutes – but the context sensitive half life is only 7-8 minutes. An infusion or repeat boluses will therefor be needed for longer procedures. The organ independent rapid metabolism makes it the ideal drug for patients with obstructive sleep apnoea or in the elderly. It might even become an option for intensive care sedation.

Alpha 2 – agonists

α2-adrenoceptor agonists have several benefits perioperatively including decreased sympathetic tone, diminished neuroendocrine & haemodynamic response to surgery, decreased anaesthetic & opioid requirement, sedation and analgesia. Although it shares some applications with Benzodiazepines, it has some added benefits and a more favourable side effect profile.

The α2-adrenoceptor

Present in the peripheral and central nervous system and in various organs such as the liver, pancreas and kidneys. It is a G_1-Protein coupled receptor that inhibits adenylate cyclase that leads to decreased cAMP formation. Modulation of ion channel activity causes hyperpolarisation of the cell

membrane. Efflux of K+ hyperpolarizes the excitable membrane and suppresses neuronal firing. A second mechanism of action is a G_0-Protein alteration of N-type voltage gated Ca_{2+} channels that decrease Ca_{2+} entry into the cell and decreases the secretion of neurotransmitters.

Action of α2-adrenoceptor agonists

The effects mediated through α2 agonism vary greatly and depend on the site of the receptor. Receptors occur presynaptically as well as postsynaptically. Effects include:

Postsynaptic inhibition in central nervous system
Hypotension & bradycardia due to lower sympathetic tone

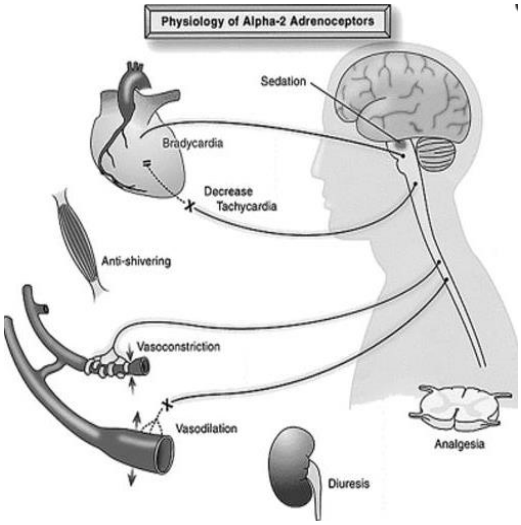
Presynaptic inhibition in Locus Coeruleus
Sedation due to decreased noradrenergic driven vigilance
Analgesia due to action in descending medullospinal noradrenergic pathway

Inhibits neuronal firing in spinal cord
Decreased firing of nociceptive neurons in substantia gelatinosa

Other areas
Decreased salivation & GIT secretion

Contraction of smooth muscle
Decreased renin release & increased GFR
Decreased insulin release

Muscle
Decreased shivering



- **Cardiovascular system**

α2-agonists display a biphasic response with an initial increase in blood pressure do to direct smooth muscle contraction with a reflex bradycardia.
This response fades after the initial bolus dose and the patient settles with lower blood pressure and a slower heart rate from baseline due to decreased sympathetic outflow from the central nervous system. The decreased sympathetic outflow is beneficial in patients with ischaemic heart disease but must be used cautiously in patients with fixed cardiac output states and heartblock.

- **Respiratory system**

α2-agonists causes no direct respiratory depression and airway tone is preserved. Increased CO_2 and slower respiratory rates are comparable with the breathing pattern during normal sleep. Oxygen consumption is reduced secondary to decreased sympathetic outflow and analgesia. Although α2-agonists do not increase respiratory depression from other respiratory depressants such as opioids, it must be used with caution when used together as the sleep induced effect is additive. There is no synergism with other sedatives.

- **Sleep Physiology**

EEG analysis of sleep spindles show that Dexmedetomidine produces a state closely resembling physiological sleep in humans, which support earlier experimental evidence for activation of normal non-rapid eye movement sleep-promoting pathways.

Dexmedetomidine

An imidazole compound that is the active dextroisomer of Medetomidine. $\alpha 2$: $\alpha 1$ selectivity of 1620:1, and more specifically for the $\alpha 2_A$ receptor, confers the benefit of improved sedation and analgesia, and less cardiovascular instability compared to Clonidine with an $\alpha 2$: $\alpha 1$ selectivity ratio of 220:1.

Dexmedetomidine is 94% percent protein bound and there is no significant displacement with other common anaesthetic drugs. It is metabolized via glucuronidation and hydroxylation in the liver by CYP2D6 enzyme activity and excreted in the urine. Virtually no unchanged drug is excreted compared to Clonidine where 50-60% of the drug is excreted unchanged. The distribution half life of Dexmedetomidine is 6 minutes with an elimination half life of 2 hours. This makes Dexmedetomidine suitable drug for short sedation procedures and infusions. Clonidine has a much longer elimination half life of 12-15 hours that makes 12 hourly oral administration a viable alternative.

Nasal administration of Dexmedetomidine is becoming popular for sedation especially for children. The bioavailability after nasal administration is 72-92% if excessive nasal secretions are cleared before administration. The onset of action is comparable with nasal Midazolam but there is less nasal irritation with administration, better levels of sedation, no paradoxical excitation and faster post procedural awakening with less confusion. Midazolam adds the quality of slight anterograde amnesia.

The pharmacokinetics of Dexmedetomidine does not really change with advancing age compared to Benzodiazepines.

Both Dexmedetomidine and Clonidine lowers the opioid requirement, prevents the sympathetic effects of opioid withdrawal, and patients do not develop tolerance. This makes it an ideal drug for ICU sedation. Patients do not experience withdrawal symptoms after the cessation of prolonged administration but may experience some rebound elevation in blood pressure that is more pronounced after Clonidine infusion. Clonidine infusions are not routinely used in ICU due to more haemodynamic instability. 12 hourly oral Clonidine are often used as an alternative to aid with sedation and analgesia instead of Dexmedetomidine infusions due to the much lower cost.

Pharmacogenomics

Pharmacogenomics in anaesthesia will become clinically more important as awareness and research into this field increases.

We are already aware of vast individual variability in patient response to sedative drugs. Paradoxical excitatory reactions to benzodiazepines could be due to mutations or variations in $GABA_A$ receptor subunits.

Variability in the vasoconstriction response to Dexmedetomidine is known to be due to a homozygous variation at the α_{2BD} allele.

We know that several different alleles exist for the cytochrome P450 CYP2D6 enzyme, which is also involved in the glucuronidation of Dexmedetomidine. This could explain the dose effect variation amongst patients.

References

1. South African Society of Anaesthesia Sedation Guidelines 2015
2. Pharmacology. Eighth Edition. Rang and Dale's. J Ritter, R Flower, G Henderson & H Rang
3. Remimazolam: The future of it's sedative potential. BG Gandra, PM Singh. Saudi J Anaes 2014 Jul-Sep 8(3) 388-391
4. Pharmoco-genomicsand anaesthesia: Mysteries, correlations and facts. SB Bhaskar et al. Indian J Anaes 2013 Jul-Aug 57(4) 336-337
5. Electroencephalogram spindle activity during dexmedetomidine sedation and physiological sleep. E Huupponen et al; Acta Anaesthesiol Scand 2008 Feb;52(2):289-94.
6. Comparison of clonidine and dexmedetomidine for short-term sedation of intensive care unit patients. U Srivastava et al; Indian J Crit Care Med 2014 Jul; 18(7): 431-436

Spinal Cord Monitoring

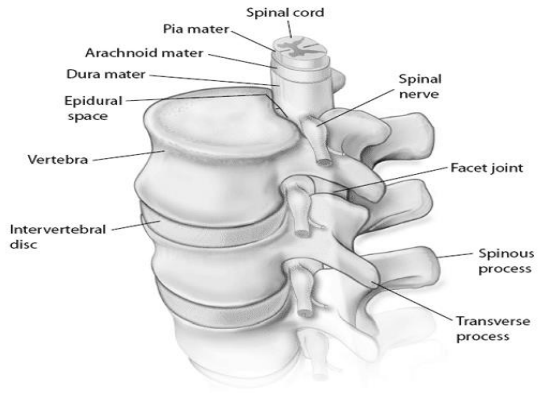
Dr Noshina Khan

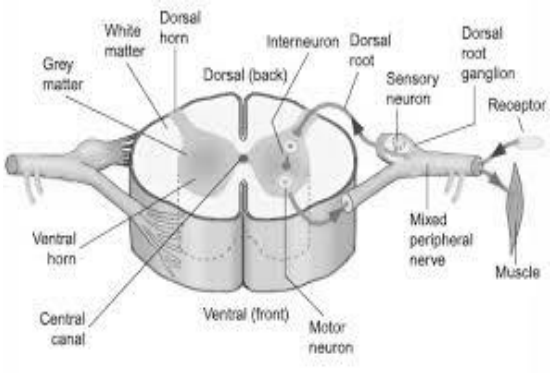
UCT Dept. of Anaesthesia & Perioperative Medicine

Neurophysiological monitoring is a continually evolving field that aims to localize and monitor neural structures according to their functional basis and ultimately preserve their structural integrity. During spinal surgery, several structures are placed at risk for potential injury, including the spinal cord, nerve roots, lumbar plexus, and all relevant vascular supply to these elements. Several electrophysiological modalities are currently available for monitoring various aspects of the central and peripheral nervous system, each offering a unique set of benefits, limitations

Please become familiar with basic anatomy of the spine as well as nerve signal transduction, from appropriate texts and sources. The aim of my notes is clear up concepts in understanding of neuro-monitoring as opposed to being a complete review, given length constraints.

Basic Anatomy Of The Bony Spine And Transverse Section Through The Spinal Cord





Anatomy of the Spinal Tracts

- Ascending tracts
- Descending tracts

Ascending Tracts

The **Ascending Tracts** refer to the neural pathways by which sensory information from the peripheral nerves is transmitted to the cerebral cortex. (Somatosensory Pathways)
Functionally, the ascending tracts can be divided into the type of information they transmit

Spinothalamic tracts
- **Lateral** (pain & temperature)
- **Anterior** (light touch & pressure)

Dorsal column medial lemniscal tract
- Deep touch & pressure
- Proprioception
- Vibration sensation

Spinocerebellar tract
- Posture & coordination

Dorsal Column

The **Dorsal Column-Medial Lemniscal Pathway** (DCML) carries the sensory modalities of fine touch, vibration and proprioception, via the *Dorsal Column* of the spinal cord to the brainstem.
In the brain stem it is transmitted through the medial lemniscus to the cerebral cortex via three orders of neurons.

First order Neurons

Convey sensory information regarding touch, proprioception or vibration from the peripheral nerves to the medulla oblongata. There are two different pathways taken by the first order neurons:

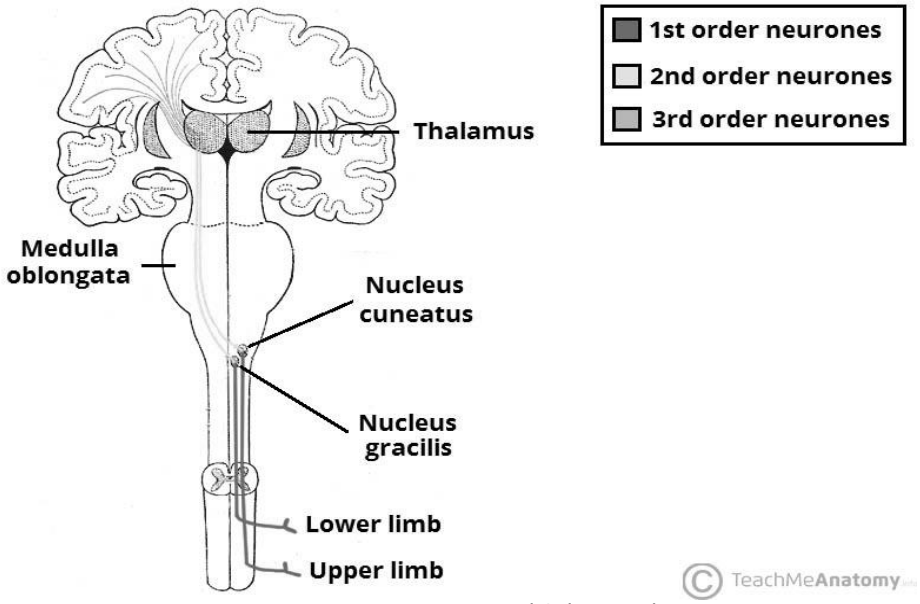
- **Signals from the upper limb** (T6 and above)
  Travel in the Fasciculus Cuneatus (the lateral part of the dorsal column).
  Then synapse in the Nucleus Cuneatus of the medulla oblongata.
- **Signals from the lower limb** (below T6) –
  Travel in the Fasciculus Gracilis (the medial part of the dorsal column).
  Synapse in the Nucleus Gracilis of the medulla oblongata.

Second order neurons

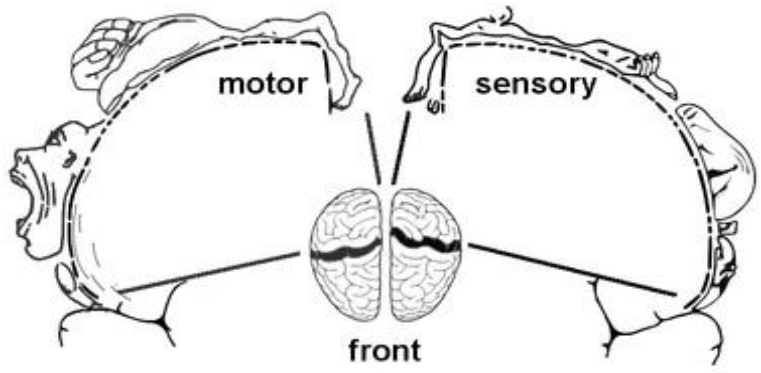
Begin in the Cuneate Nucleus or Gracilis.
The fibers receive the information from the preceding neurons, and deliver the information to the third order neurons in the thalamus.
Fibers **decussate** within the **Medulla Oblongata** (cross to the other side of the CNS), then travel in the **contralateral Medial Lemniscus** to reach the **Thalamus**.

Third order neurons

Transmit the sensory signals from the **thalamus** to the **ipsilateral Primary Sensory Cortex of the brain**. They ascend from the ventral Posterolateral Nucleus of the thalamus, travel through the internal capsule and terminate at the sensory cortex, and distribute in the Sensory Cortical Homunculus originally mapped by the neurosurgeon Wilder Penfield

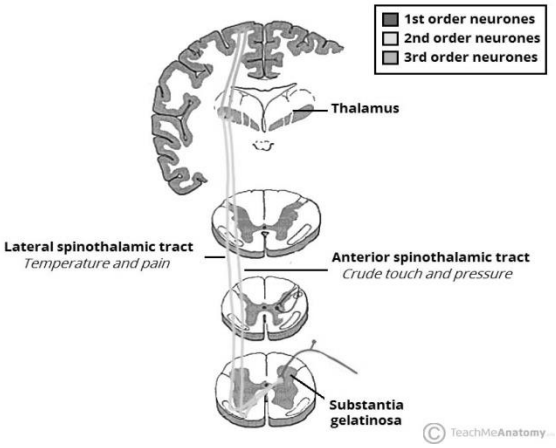


Dorsal Column Pathway



Homunculus for sensory and motor cortex

Spinothalamic Tracts



Spinocerebellar Tract
Please refer to relevant texts for more information on details of tracts not described in this text.

Descending Pathways

The **descending tracts** are the pathways by which **motor signals** are sent from the brain to LMN. The LMN then directly innervate muscles to produce contraction. The motor tracts can be functionally divided into two major groups:

- **Pyramidal tracts**
  Originate in The Cerebral Cortex, carrying motor fibres to the spinal cord and brain stem. They are responsible for the *voluntary control* of the musculature of the body and face.
- **Extrapyramidal tracts**
  Originate in the brain stem, carrying motor fibres to the spinal cord. They are responsible for the *involuntary and automatic control of all musculature*, as muscle tone, balance, posture and locomotion.

There are **no synapses within the descending pathways**.
At the termination of the descending tracts, UMN synapses with the LMN.
All the neurons within the descending motor system are upper **motor neurons**. Their cell bodies are found in the cerebral cortex and the brain stem, with their axons remaining within the CNS.

Pyramidal tracts

Functionally, these tracts can be subdivided into:

- **Corticospinal tracts** – supplies the musculature of the body.
- **Corticobulbar tracts** – supplies the musculature of the head and neck.

The Corticospinal Tracts

The corticospinal tracts begin in the cerebral cortex; from which they receive a range of inputs:

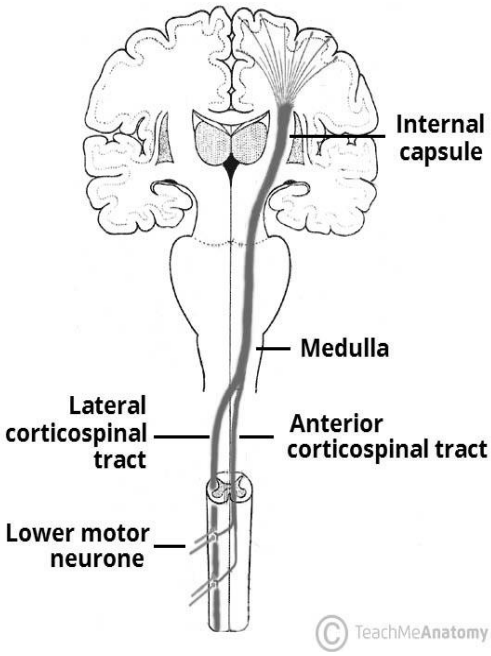
- **Primary motor cortex**
- **Premotor cortex**
- **Supplementary motor area**

They also receive nerve fibers from the **Somatosensory Area**, which play a role in regulating the activity of the ascending tracts.

After originating from the cortex, the neurons converge, and descend through the **Internal Capsule** (a white matter pathway, located between the thalamus and the basal ganglia), then pass through the **Crus Cerebri** of the midbrain, the **Pons** and into the **Medulla**.

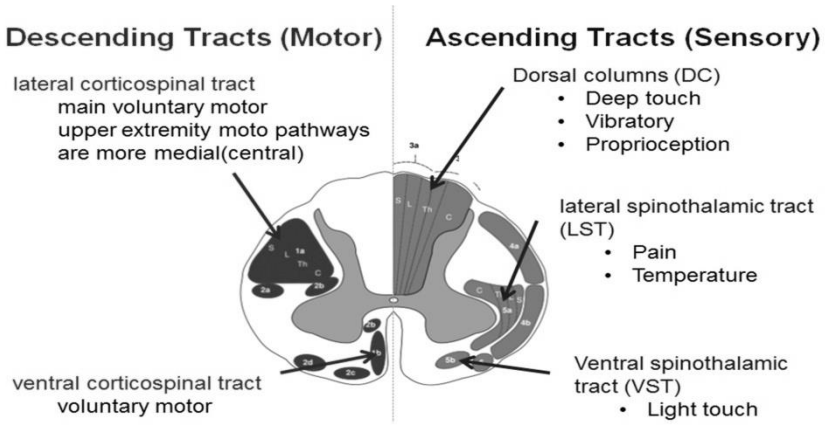
In the most inferior (caudal) part of the medulla, the tract divides into two paths. 75-90% of the fibers *decussate* (cross over to the other side of the CNS), forming the ***Lateral Corticospinal Tract.*** They then descend into the spinal cord, terminating in the ***Ventral Horn*** *(Anterior Horn)* (at all segmental levels). From the **ventral horn,** the ***LMN*** go on to supply the muscles of the body. The CST is the main channel responsible for voluntary movement and is the main determinant of signal generation during transcranial motor-evoked potential (TcMEP) monitoring.

The remaining fibers on the ipsilateral side form the **anterior corticospinal tract**, descending into the spinal cord. They then decussate and terminate in the ventral horn of the **cervical** and **upper thoracic segmental levels.**



Corticospinal Tract

The axons of LMN then leave the spinal cord with their respective anterior nerve roots becoming ***spinal nerve roots*** in the neural foramina before recombining in the respective ***nerve plexus***, forming ***peripheral nerves*** and terminating at the neuromuscular junction.
The axon of a single LMN innervates a number of muscle fibres, creating a ***motor unit***. The lower the number of muscle fibres within a motor unit, the finer the movement.
Synchronized firing of multiple LMN is vital to achieve a visible muscle contraction.
Muscle groups usually receive an input from the LMN of neighbouring spinal levels and therefore multiple nerve roots that intermingle in the cervical and brachial plexuses before forming peripheral nerves. This is termed *radicular overlap*, a redundancy that may mask injury to individual nerve roots during spinal cord monitoring.



Tran sectional View Of Spinal Cord Illustrating The Various Tracts

Blood supply and Perfusion of the spinal cord

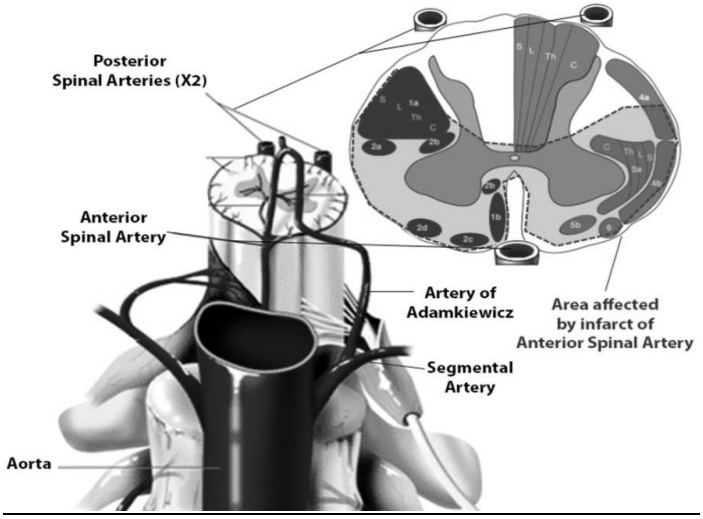
The homeostasis and function of the spinal cord and its nerve roots are primarily determined by adequate blood flow and oxygen supply. *Blood flow* is predominately distributed within the spinal cord *via three vertical arteries*.

- **Anterior Spinal Artery**
  - Supply of anterior 2/3 of spinal cord
  - (lateral corticospinal tract and ventral corticospinal tract)
- **Posterior Spinal Artery (Right & Left)**
  - Supply to posterior 1/3 of the spinal cord
  - (dorsal sensory columns)

The *anterior spinal artery (ASA) runs in the ventral median fissure* and supplies the *anterior two-thirds* of the spinal cord, also covering most of the *gray matter through end arterie*s. The *two posterior spinal arteries (PSA) run alongside the dorsolateral columns* and supply mainly *the white matter of the posterior one-third of the spinal cord*. The remaining anterolateral parts of the white matter are supplied through a ***circumferential anastomotic plexus between the ASA and PSA***. Blood flow through the ASA and PSA is primarily determined by a rich collateral network from the subclavian and iliac systems, and from aortic segmental arteries.

Spinal Cord Autoregulation, which maintains constant blood flow between wide ranges of MAP, has very similar range limits to the brain (MAP 70- 50mmHg). Below the critical lower limit of this range, blood flow becomes pressure passive. There is much inter individual variation as to this critical pressure, dictated by patient comorbidities as well as external events like surgical traction, anaesthetic agents. Consider targeting patient appropriate MAP to ensure one does not hypo perfuse the cord.

<u>Blood Supply to Spinal Cord</u>



Mechanisms of Intra Operative Injury

<u>Direct</u>

Mechanical compression
Trauma via instrumentation

<u>Indirect</u>

Disruption in O2 demand – supply balance
Disrupted autoregulation- disease, anaesthetics
Hypo perfusion
Hypotension, blood loss

Monitoring Physics

Before the introduction of IONM into clinical practice, the surgical team relied on the **Stagnara wake up test o**r ankle clonus to confirm the integrity of spinal cord during procedures potentially risky for the cord. The drawbacks of this test are that they disrupt the surgical procedure and they cannot be performed continually. Moreover, sudden movement of the patient during these tests might produce damage to the cord, and required much patient priming and cooperation. It is also inappropriate to use with developmental delay and deaf patients. Also, given the delay to wake up and actually get patient to move legs, valuable time for remedial action is indeed lost. Intraoperative neurophysiologic monitoring (IONM) is a valuable technique for assessing the nervous system. It replaces the neurologic examination when the patient is under general anaesthesia and cannot cooperate with a face-to-face examination, allowing for assessment of many neural structures including the neuromuscular junction, peripheral nerve, spinal cord, brainstem, and cortex during surgery. The monitoring modalities focused on will be those used for monitoring SPINAL CORD integrity intra operatively. The most commonly employed techniques during spinal procedures are:

Transcranial motor evoked potentials (Tc-MEPs)
Upper and lower somatosensory sensory evoked potentials (upper and lower SSEP),
Pedicle screw simulation
Spontaneous electromyography (EMG)

A number of other techniques have been used over the years that include direct spinal cord stimulation and reflex monitoring.

1. Somatosensory Sensory Evoked Potentials (SSEP)

The SSEP was the first effective means for monitoring the function of the spinal cord during surgery. It is a recorded central response to stimulation of a peripheral sensory nerve. SSEP attempts to ascertain integrity of the sensory dorsal column pathway of the spinal cord. Peripheral receptors are stimulated and the information relayed via neurons whose stroma are located in the dorsal root ganglia at all spinal levels. (It should be noted that SSEPs do not measure the slower conducting fibres of the spinothalamic pathway.) Axons from these first-order neurons project to the spinal cord via the medial root entry zone, giving rise to the Fasciculi Gracilis and Cuneatus, which subsequently carry sensory information from the lower and upper extremities, respectively. ***The first synapse in this pathway occurs in the lower medulla after these tracts ascend via the dorsal columns*** in the spinal cord. (i.e. no synapse along the actual cord). Following a decussation (medullary level), the medial lemniscus is formed; it then ascends to the thalamus and ultimately relays sensory information to the Primary Somatosensory Cortex (Brodmann 3,1,2)

In the upper extremity, peripheral nerve stimulus is typically applied to

- Median nerve (C-6, C-7, C-8, and T-1 roots) and
- Ulnar nerve (C-8 and T-1), usually in the wrist but more proximal stimulations are possible.

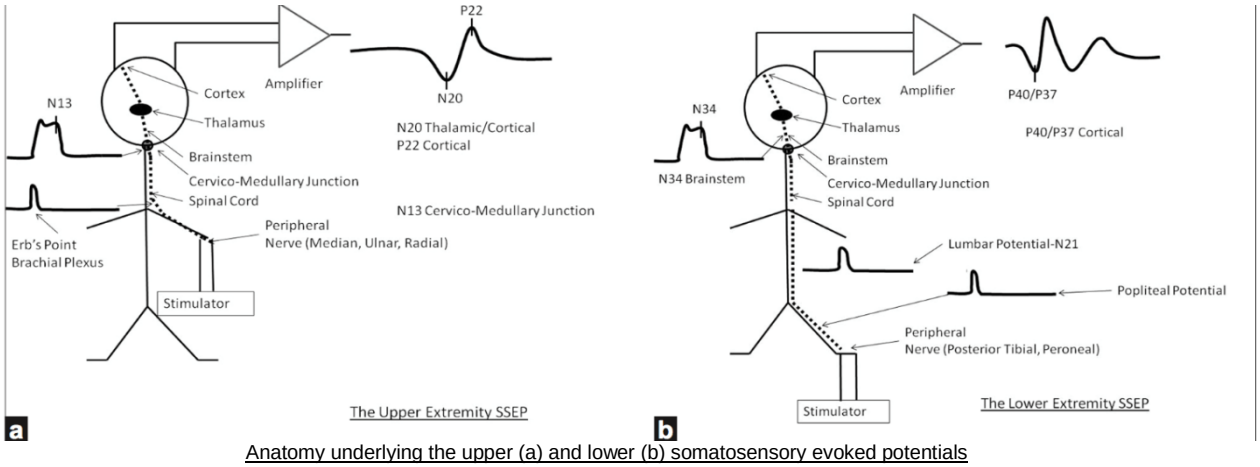
In the lower extremity, the typical site of stimulation

- Posterior tibial nerve (L-4, L-5, S-1, and S-2) and
- Peroneal nerve (L-4, L-5, and S-1).
  Alternate stimulation sites include the peroneal nerve at the fibular head and the tibial nerve in the popliteal fossa.

Signal transduction most often recorded at the sensory cortex, however the signal may be measured at various points along its path, (like at the brachial plus, spinal cord, cervico-medullary junction). The action potential is predicted to arrive at specified points along its conduction path at predicted time intervals (**latency**) as well as with a predicted intensity (**amplitude**).

Anatomy Of The Somatosensory Evoked Potential (for interest)

Electrical stimulation in the extremities produces major positive and negative deflections as signals that ascend via the somatosensory pathway. Most often, a *negative potential* is measured at the *scalp* corresponding to *the upper extremities at 20 milliseconds (N20),* and a *positive potential* is measured at the scalp corresponding to the *lower extremities at 37 milliseconds (P37*). As the impulses from the lower extremity enter the cauda equina, a lumbar potential (N21) is generated. A peripheral response recorded at the level of the brachial plexus (for the upper extremities) or the popliteal fossa (for the lower extremities) can be performed to ascertain adequacy of stimulation. These peripheral responses can also help to detect peripheral limb ischemia or nerve compression, as a result of poor positioning.



Anatomy underlying the upper (a) and lower (b) somatosensory evoked potentials

Advantages

- SSEP's have a simpler wave form to quantify compared to MEP's
- Less affected by type of anesthetic agents as a result of fewer synapses along the spinal cord as compared to MEPs – within reason.

Problems with SSEP

- Both the popliteal and lumbar potentials can be difficult to record, especially in patients who are **overweight.**
- The amplitude of a single compound action potential is not sufficiently high to be detectable in the noisy background of an electroencephalogram (EEG), opposed to MEP waveforms. **SSEP require averaging of several sweeps of signals**- (i.e. generated by multiple stimulation sweeps.) **This takes up to 3-5 min**. The delay in information relay may cause surgical disruptions, and makes for difficulty with real time interpretation and subsequent remedial action.
- The **signal-to-noise ratio** (SNR) varies at different points along the sensory pathway and in different regions of the cerebral cortex. Under general anesthesia, the frontal cortex generates more EEG noise and so a lower SNR compared with parietal or posterior cortices interestingly!
- SSEP's are **not reliable for monitoring the integrity of the anterior motor pathways**. (SC supplied by anterior spinal artery). Reports exist in literature of an ischemic injury leading to paralysis despite normal SEP monitoring during surgery, because disruption of anterior spinal artery blood supply has affected only the motor pathway.
- Although SSEP signals are good basic indicators of spinal cord function, less information is provided regarding actual nerve root function. SSEP's are a composite of *summated neural signals* that enter the spinal cord through *multiple segments*. And so it is possible that a segmental injury is missed due to cross innervation. In addition, due to central amplification, it is possible for SSEPs to remain completely normal in the face of a nerve root injury.
- Factors that potentially affect the SSEP amplitude include halogenated agents, nitrous oxide, hypothermia, hypotension, and electrical interference. (Albeit SSEP's are less sensitive to anesthetic agents compared to MEP's, increased latency and depressed amplitude is a dose related effect of halogenated agents.)

Criteria for change in somatosensory evoked potentials (interest)

Interpretation of what represents a significant change remains controversial because of intra observer variability and because the optimal criteria for determining when there is a significant change are very dependent on the type of procedure. It is generally accepted that injury to the large fibre dorsal column pathways is typically expected when there is a ***50% reduction in signal amplitude or 10% prolongation of latency from baseline, especially if the changes persist beyond two averaged cycles***

2. Motor Evoked Potentials (MEP)

Prior to the widespread use of MEP monitoring, the only way to assess corticospinal tract integrity and motor function during surgery was the Stagnara wake-up test. Function of MEP is to monitor integrity of lateral and ventral corticospinal tract of the spinal cord
In MEP monitoring, **Motor Pathways** (*Lateral and Ventral Corticospinal Tracts*) are **assessed directly**, solving a major limitation of SSEPs in determining the integrity of motor neurons. In a sense, MEPs measure the integrity of the motor neuron output. A **stimulus** is initiated at a central site, usually via **Motor Cortex**. Stimulation of the motor cortex thru the scalp is called transcranial motor evoked potential (TcMEP). Cortical motor evoked potentials are when the motor cortex is stimulated directly (i.e.: thru a craniotomy).

Stimulation can also be performed in the spinal cord directly, with recording electrodes either in the nerve or the muscles. Although this technique offers the advantage of being less sensitive to anaesthetic agents, responses obtained via direct spinal cord stimulation are less likely to represent motor function, but rather antidromic sensory responses.

A **signal recording** is made of the **evoked responses** occurring distally, at the level of the muscle (**Compound Muscle Action Potential**), nerve (neurogenic MEP) or spinal cord (direct corticospinal wave **D-wave** recording.

Transcranial Motor Evoked Potentials

Tc-MEPs have been used to perform intraoperative monitoring for more than 20 years. The Tc-MEP involves applying a train of high voltage stimuli to electrodes on the surface of the head to activate motor pathways and produce either a motor contraction (muscle MEP) or a nerve action potential (D-wave) that can be recorded.

In a ***normal awake patient,*** electrical stimulation of the cortex/subcortical white matter with a *single* electrical pulse produces a number of responses that can be recorded by an *epidural electrode placed over the upper thoracic spinal cord.* The first of these waves is called the **direct or D-wave** and the **succeeding waves are termed indirect or I-waves**. The D-wave is the “impulse produced” (orthodromic) nerve action potential that results from stimulating white matter directly. It involves *no synaptic activity*. The *I-waves* represent the discharges produced by the *nearby cortical neurons* that were excited by the same stimulus. *These require synaptic activity* and are hence strongly suppressed with volatile anaesthesia that depresses signalling at a synapse level as opposed to interference with axonal transduction. (one of the mechanisms)

This is also important because of the characteristics of the ***anterior horn cell***, which is the final common pathway for all motor responses. These cells respond optimally to *multiple sequential*, *precisely timed stimuli* compared to a single stimulus, to ensure firing. Thus, an anterior horn cell will fire easily in response to a train of stimuli, but will not always readily respond to just a single stimulus. So under anaesthesia, with disarrangement of synaptic signalling, one is more likely to get a response from the AHC via initiating a pulsed train of stimuli, as opposed to a single stimulation as with the awake patient.

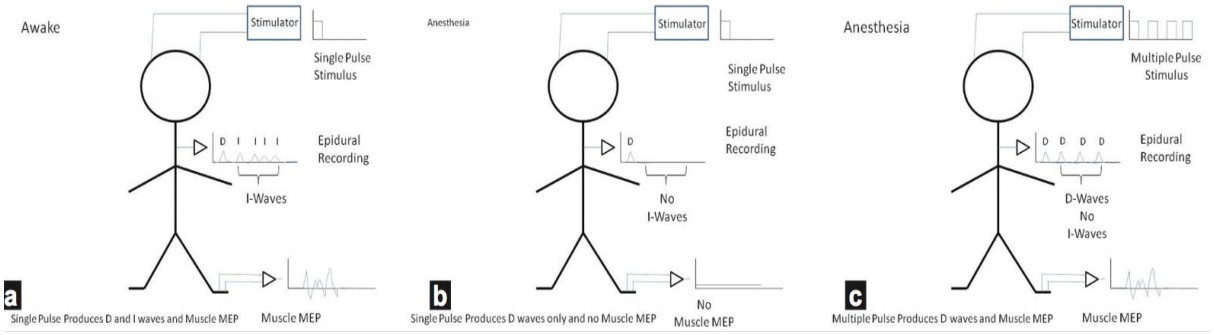


Illustration of the responses to a single-pulse transcranial electrical stimulation in the (a) awake and (b) anesthetized patients. (c) Shows the effect of multi pulse stimulation in the anesthetized patient

Stimulus Parameters and Transcranial Motor Evoked Potentials (interest)

Stimulation trains increase temporal summation at the alpha motor neuron. The train stimuli used during TC-MEP range in amplitude from about 75 to 900 V, with maximal currents up to 0.9 A. The stimulation voltage and current required is markedly dependent on the type of electrode used. The highest current levels are required if **EEG cup electrodes** are used. Lower thresholds are seen **with subdermal needle electrodes** and **corkscrew electrodes**. Electrodes are placed over specific positions of the head. The duration of each pulse is between 50 and 500 msec. The longer duration pulses are associated with a lower threshold. The number of pulses ranges between 3 and 12, with the frequency of the pulses at 150–500 Hz. Latencies of 20 ms in the hand and 45 ms in the foot are typically observed, depending on various factors such as the underlying pathological condition, the patient's height, and body temperature.

Response- Muscle or D wave

The muscle MEP is the most commonly used Tc-MEP. Recordings are of high amplitude and can be obtained with a single sweep. Thus, they can provide the surgeon with nearly instantaneous information, unlike the SSEP that requires prolonged averaging. The problem with the muscle MEP is that the waveform is complex. Other limitations include significant intrinsic variability, gradual signal drift toward higher thresholds over time while under anaesthesia. When compared to SSEP, obtaining an MEP signal is significantly more challenging with a 93% versus 66% respective success rate for lower extremities. MEPs are sensitive to underlying neuromuscular disorders and drugs commonly used during general anaesthesia. Diabetes, hypertension and anaesthetic techniques are the most important risk factors associated with failure to obtain lower extremity MEP signals.

D-wave

Because the D-wave has a simple morphology and is insensitive to anaesthesia, criteria for interpretation are much simpler than for the muscle MEP. It is generally considered that 50% decline in amplitude is an appropriate alert criterion. Disappearance of the D-wave is generally associated with a significant neurologic deficit. Although this technique is very powerful and responses can be obtained even in the presence of neuromuscular blockade with a single stimulus, it does have some limitations. First, it must be recorded with an epidural electrode. Second, it is difficult to record below the mid-thoracic region. Electrodes migrate. Finally, the changes in the configuration of the spinal canal during surgery may cause changes in the D-wave, and thereby limit its use in these cases. Despite these limitations, D-wave recording remains particularly important during spinal cord tumour surgery for upper cord.

TcMEP and Anaesthesia

Under general anaesthesia, a single stimulus may not be effective in producing a response, since the I-waves are diminished and the anterior horn cell sees only the single D-wave. In addition, during general anaesthesia, there is a reduction in spontaneous activity of the interneurons of the spinal cord, reducing the overall level of excitation reaching the anterior horn cell. These problems are overcome by using multi pulse stimulation as opposed to single pulse (as well as anaesthetic modifications). The depression of intrinsic spinal cord activity is greatest with the halogenated anaesthetic agents and nitrous oxide. Since the D-wave is purely a nerve action potential and does not involve synaptic activity, it is relatively insensitive to the effects of anaesthesia.

Interpretation of the MEP recording (interest)

Currently accepted methods of interpretation for TcMEP responses:

The All-Or-Nothing Criterion: Complete disappearance of an MEP signal reduces false positives but misses subtle but significant events. ASNM recommends the all-or-nothing criterion is considered a major criterion and mandates corrective surgical action in orthopaedic spine surgery. In a study of monitoring during surgery for intramedullary spinal cord tumours, loss of the muscle MEP without more than a 50% change in the D-wave was associated only with transient neurologic deficits. The problem is that in most spinal surgeries, other than spinal cord tumours, D-wave recording is difficult and limited.
The Amplitude Criterion*:* proposed that a reduction in amplitude of 50% or more should be considered as significant. The problem with this is that there is quite a bit of natural variability in the muscle MEP, which may increase the false-positive and false-negative rate. Some prefer a reduction

of at least 80% amplitude in 6 muscle groups.
The Threshold Criterion: This criterion was based upon the fact that the stimulus threshold for obtaining a muscle MEP increases when there is damage to the corticospinal tract. Increases of more than 100 V in the threshold for a sustained period of time in order to obtain a muscle MEP are considered an early sign of injury. The difficulty with this criterion is that thresholds generally increase gradually during surgery and are significantly influenced by even small changes in anaesthesia. ASNM recommend threshold and amplitude a moderate criterion
The Morphology Criterion: A sustained change from polyphasic to biphasic CMAP waveform morphology appears to correlate with significant events in spinal cord tumour surgery.

The effect of pre-operative damage to the motor pathways and the transcranial motor evoked potentials Damage to motor pathways prior to surgery has impact on generation of the muscle MEP. If there is injury pre-operatively, even if the patient has good strength pre-operatively, the MEPs may be difficult to obtain. This is because activation of the AHC requires a highly synchronized sequence of inputs that can easily be desynchronized by a minor disruption of conduction, as previously stated.

Complications of TcMEP

Electrical shock. Because of the high voltage and high current delivered with TcMEP, there is the risk of shock to staff that come into contact with the electrodes.
Regional **tissue burns/ injury** may occur in the area of the scalp electrode
Local areas of **hair loss** have been reported
Bite injuries. TcMEP stimulation can cause direct stimulation of trigeminal nerve, hence jaw contractions. The bite injury of tongue or lips can occur. Mandibular contractions may also result in patient biting through the endotracheal tube creating an emergency. As well as patient migration from the gel head support, thereby possibly creating **malposition** with subsequent **pressure related injuries**. Ensure frequent pressure checks intra op.
Movement injuries. The patient may move during the elicitation of TcMEP's, injuring a vital structure if it is jolted or torn away.
Seizures. The possibility that brain stimulation could provoke a seizure is always there; albeit small, especially with pulse train stimulation. It is however still recommended to avoid TcMEP in those at risk of seizures, or at least to do so very infrequently.
Cardiovascular Complications. Cardiac arrhythmias or blood pressure changes have been observed but the exact relationship is unclear. It is thought to be due to current penetration of the hypothalamus or brainstem. Transcranial electrical stimulation artefact may mimic cardiac arrhythmias.
In presence of implanted defibrillator, it is prudent not to perform the study unless there is very high risk of motor injury. Consultation in advance with a cardiologist is suggested.
Patients with a pacemaker without defibrillator, the risk of damage or aberrant firing of the pacemaker is low, though consultation with the cardiologist recommended.

Contra indications to TcMEP

Relative contraindications to TcMEP stimulation include epilepsy, convexity skull defect, cortical lesions, raised intracranial pressure, cardiac disease, pro-convulsant medications or anaesthetics, intra-cranial electrodes, vascular clips or shunts and cardiac pacemakers or other implanted biomedical devices. In an infant with open fontanelle or open suture do not use spiral needle electrode.

3. Spontaneous EMG

The recording of spontaneous EMG activity from a muscle provides information on the state of the peripheral nerves that innervate that muscle. Compression or stretch of a nerve as well as hypothermia and ischemia produce depolarization of the axons resulting in the appearance of spontaneous action potentials. These action potentials subsequently produce contractions of muscle fibres that can be recorded by electrodes placed in the muscle. (Further details beyond scope of talk)

4. Triggered Electromyography

For instrumented spinal fusions, electrical stimulation facilitates proper screw placement. The most common use is to determine whether a screw that has already been placed is properly located. The basic principles that if the screw is electrically close to one of the nerve roots, then electrically stimulating the screw will activate the nearby nerve root at a low current level.

5. Direct Spinal Cord Stimulation

Spinal cord stimulation techniques involved stimulating the spinal cord either directly or through a long intraosseous electrode. Recordings were made from peripheral nerves. This technique was originally proposed as a means of monitoring motor function that could be achieved even with total paralysis since only the peripheral nerve action potential was recorded, and not the muscle response. However, collision studies have demonstrated that the responses are mainly the result of conduction along large fibre somatosensory pathways similar to those used by the SSEP, and hence this technique does not provide additional information to that already conveyed by the SSEP and MEP.

ANESTHESIA AND SSEP & MEP

The effects of anaesthetic agents are thus the result of direct inhibition of synaptic pathways or by indirect effect on pathways by altering the balance between inhibitory and excitatory signalling mechanisms. Inhalational anaesthetic agents essentially reduce synaptic transduction, while Ketamine and Etomidate belong to the latter category. Thus, while all other anaesthetic agents depress amplitude and increase the latency, Ketamine and Etomidate increase the amplitude perhaps by attenuating inhibition. The electrophysiological responses that rely excessively on synaptic function will be most susceptible to anaesthetics. The DC has fewer synapses along the spinal cord pathway than the motor tract, hence the DC albeit affected is less affected by volatile agents. This also explains why cortical responses are more sensitive to anaesthesia than those generated in the periphery such as brainstem or spinal cord, as fewer synapses are involved. Often it's found that a drug may affect SSEP's measured at a cortical level but not when response is sought below this. For simplicity I am only going to focus on signalling via the cortex- cortical SSEP and TcMEP.

Anaesthetic Agents affecting Evoked Potentials

Inhalational

Thought to affect signal transmission by altering synaptic function. All halogenated inhalational agents produce a dose-related increase in latency and reduction in the amplitude of the cortically recorded SSEPs. Motor evoked potentials single pulse stimulation TcMEP's recorded at muscle level are essentially abolished (unrecordable) in the presence of inhalational agents. This effect of inhalational agents is likely the result of depression of synaptic transmission either in the anterior horn cell synapses on motor neurons or in the cortex on the internuncial synapses with a resulting loss of I waves.

Nitrous oxide

Nitrous oxide reduces SSEP amplitude and increases latency when used alone or when combined with halogenated inhalational agents or opioids. It potentiates the effects of other agents, including high dose Propofol. N_2O has similar effect on TcMEP as the halogenated gases.

Opioids

The effects of the opioid analgesics (alfentanil, fentanyl, remifentanil, and sufentanil) on SSEPs and MEPs are less than with inhalational agents, making opioids important components of anaesthetic planning for evoked potential monitoring. There is some depression of amplitude and an increase of latency in both SSEP and MEP's but signal recording is still possible. Most depression occurs after a massive bolus administration.

Ketamine

Ketamine can enhance cortical SSEP amplitude and MEP amplitude.

Barbiturates

Thiopentone remains a popular drug for induction of anaesthesia, although transient decreases in amplitude and increases in latency of SSEPs occur immediately after induction. Barbiturates are not used commonly during recording of MEPs because the CMAP responses are unusually sensitive to the effect appears to be quite prolonged. In one study, the induction bolus eliminated the CMAP from

the MEP for a period of 45 to 60 minutes!

Midazolam

Mild depression of cortical SSEP's occurs at doses of 0.2mg/kg. Doses as low as0.05 mg/kg of Midazolam produces prolonged, marked depression of MEPs, suggesting that it also may be a poor induction agent for MEP recording.

Etomidate

This amplitude increase appears coincident with myoclonus seen after administration of the drug, suggesting a heightened cortical excitability. Studies with MEPs have suggested that Etomidate is an excellent agent for induction and monitoring of CMAP responses. Of several intravenous agents studied, Etomidate had the least degree of amplitude depression after induction, but one has to consider its other clinical effects as well.

Propofol

Propofol induction produces amplitude depression in cortical SSEPs with rapid recovery after termination of infusion. MEPs have demonstrated a depressant effect on response amplitude, consistent with a cortical effect. Higher doses of Propofol infusion have been linked to poor signal production. (Propofol does not appear to enhance cortical responses; rather rapid metabolism allows rapid adjustment of the depth of anaesthesia and effects on evoked responses.)

Dexmedetomidine

Dexmedetomidine as does not change SSEPs or MEPs response during complex spinal surgery by any clinical significant amount, when given at appropriate doses. It is a useful adjunct that I personally use regularly.

Muscle Relaxants

Because muscle relaxants have their major site of action at the neuromuscular junction, neuromuscular blockade will prevent recording of the CMAP during myogenic MEP monitoring. In some instances, it may be useful. Partial neuromuscular blockade (only 2-3 twitches on TOF) has the benefit of reducing a substantial portion of patient movement, which accompanies the testing, and facilitates surgical procedures. Muscle relaxation may be necessary to reduce the noise in recording electrodes (e.g. SSEPs recorded near neck or facial muscles). When using a neuromuscular blockade, tight control of the blockade is necessary so that excessive blockade does not eliminate the ability to record MEPs, thereby mimicking neural injury. For this reason, most clinicians use drug infusions, some with closed-loop control systems, to monitor the twitch and to control the infusion. In particular, the avoidance of drug boluses is important to prevent wide changes in relaxation, especially excessive degrees of block.

Although recording of myogenic MEP responses is possible with partial neuromuscular blockade, the amplitude of the CMAP will be reduced by the blockade. This technique is not popular.

Anaesthetic agent	SSEP	MEP
Halothane (0.5-1 MAC)*	↓A ↑ L	++
Isoflurane (0.5-1 MAC)*	↓A ↑ L	++
Sevoflurane (1.5 MAC)*	↓A ↑ L	++
Desflurane (1.5 MAC)*	↓A ↑ L	++
Nitrous oxide (60-70%)	↓A –L	++
Barbiturates	↓A ↑ L	++
Propofol	↓A	++
Ketamine	↑A	+
Etomidate	↑A	+
Opioids	↓A ↑ L	–
Benzodiazepines	–	+++
Dexmedetomidine	–	+
Neuromuscular blockers	–	+++

MAC = Minimum alveolar concentration, SSEP = Somatosensory evoked potential, MEP = Motor evoked potential, A = Amplitude, L = Latency, – = Negligible or no effect, + = minimal effect, ++ = significant effect, +++ = Profound effect, *without nitrous oxide

Summary of effects of Anaesthetic Agents on Evoked potentials

Non Anaesthetic Factors Influencing MEP/SEP

Blood Flow and Blood Pressure

There is a threshold relationship between regional cerebral blood flow and cortical evoked response. The cortical SSEPs responses are unaffected until blood flow is decreased to 20 ml/min/100 g. Between 15 and 20 ml/min/100 g blood flow, the SSEPs are affected and finally lost. Local factors may also cause regional ischemia. In spinal surgery, the effects of hypotension may be aggravated by spinal distraction, such that an acceptable limit of systemic hypotension cannot be determined without monitoring.

Hypoxia

Hypoxemia can also cause evoked potential deterioration before other clinical parameters are changed. Mild hypoxia ($PETO_2$ of 48 mmHg) does not influence human SSEPs. Early response to ischemia or hypoxia can manifest as a transient increase in SSEP amplitude. MEP's are lost in the face of hypoxia.

Intracranial Pressure

Increased intra-cranial pressure leads to reduction of cortical response, and so decreases amplitude and increases in latency of cortically generated SSEP's. MEPs- there is gradual increase in onset of response (latency) until it can no longer be produced.

Blood Rheology

Because changes in haematocrit can alter both oxygen carrying capacity and blood viscosity, the maximum oxygen delivery is often thought to occur in a midrange haematocrit (30 to 32%). Evoked response changes with haematocrit are consistent with this optimum range. Increase in amplitude with mild anaemia, while an increase in latency at haematocrits as low as10 to 15% are found with SSEP and further latency changes and amplitude reductions at haematocrits less than 10% MEP's have a similar response.

Ventilation

Alterations in carbon dioxide are known to alter spinal cord and cortical blood flow.
SEP and MEP are altered at critically low CO^2 as this is thought to be due to a disruption of perfusion.

Temperature

Hypothermia to 35°C decreased central and peripheral nerve conduction velocities, SSEP latency and central conduction time increased by 10-20%. These changes in SSEPs induced by hypothermia return to baseline after 30 min of re-warming. While hypothermia induced alterations of SSEP latencies are well defined, amplitude behaves in an unpredictable manner.
Hypothermia may also change the plasma concentration of anaesthetics and neuromuscular blockade and therefore, influence the MEPs. A temperature reduction of 3°C increased blood Propanol concentration by 30% during constant rate of infusion
Regional temperature changes can also alter evoked responses that would not be otherwise predicted based on unchanged core temperature. Irrigation of spinal cord, brainstem, etc., with cold saline causes routine alterations in evoked responses. For the same reason, limb cooling (from cold infusion of fluids) can change the SSEP originating from stimulation to a nerve from that extremity.
Hypothermia also increases stimulation threshold. Similar though more perverse findings occur with hypothermia and MEP's

General Notes

Plan your anaesthetic technique around the surgery performed, structures at risk, and so the most suitable monitoring for use. Discussion with the surgeon is essential. As only after one decides which signalling pathways are involved, and the level of such response and signalling that is to be used, can we decide what anaesthetic method is most appropriate. Some concepts however are universal.

Always create a stable anaesthetic environment prior to recording a baseline signal and do not alter anaesthetic technique or depth throughout the procedure, as this will cause signals to differ from the

baseline. Haemodynamic stability and ensuring adequate spinal cord perfusion should be a priority, as is maintaining normothermia throughout the procedure. A bolus administration of an intravenous anaesthetic will result in complete signal loss during a critical phase of surgery. Try and choose an anaesthetic agent with rapid onset of action, and offset and minimal effect on evoked responses. In the case of MEP's, this would mean a TIVA with Propofol in all likelihood. Even this is not benign. Remember that rapid large bolus of the agent will cause loss of signal. Communicate this to the surgeon. Give consideration to adjuncts like Ketamine, that not only assist with analgesia, as well as decrease the Propofol requirements, but also to some degree improves amplitude of signalling.

References

1. Anaesthetic considerations for evoked potentials monitoring, P K Bithal, Journal of Neuroanaesthesiology and Critical Care (1) Jan-Apr 2014
2. A review of intraoperative monitoring for spinal surgery, M. Stecker, Supplement to Surgical Neurology International, SNI: Spine (2012)(3)(3)
3. Neurosurgical Focus, Intraoperative Neurophysiological Monitoring during Spine Surgery: A Review
4. Andres A. Gonzalez; D Jeyanandarajan; Chris Hansen; G Zada; P.C. Hsieh, Neurosurg Focus. 2009(27)(4):E6
5. Neurophysiological monitoring and spinal cord integrity, F. Rabai, R.Sessions, C.N. Seubert, Best Practice & Research Clinical Anaesthesiology 30 (2016) 53e68

Online resources:

Google
Wikipedia
You tube
Anaesthesia.uk
Ortho bullets

The Splanchnic Circulation

Dr Dee Batty

UCT Dept of Anaesthesia & Perioperative Medicine

Keypoints

- The splanchnic region participates in the regulation of circulating blood volume and systemic blood pressure. It is auto-regulated like the cerebrovascular and renal circulations.
- It provides a reservoir of blood in times of circulatory stress. Major reduction of splanchnic blood volume and flow defends the perfusion of the brain and the heart in acute hypovolaemia, but prolonged hypoperfusion of the splanchnic region will lead to hypoxic tissue injury.
- Regulation of the splanchnic circulation is highly complex and variable depending on systemic physiological status, the presence of pathology and therapeutic interventions employed.
- The effect of therapeutic interventions themselves will also be context sensitive. The splanchnic circulation is in close interaction with the systemic haemodynamics under normal conditions. In low blood flow situations the response of the splanchnic circulation depends on concurrent haemodynamic factors. In the intensive care patient at risk of multiple organ failure, there is a complex and poorly understood interaction between the splanchnic blood flow, metabolic demands of the tissues and the mediators of inflammation and vasoregulation
- Most of our knowledge of the splanchnic circulation is derived from animal studies and indirect measurement.

What we do not know about the splanchnic circulation

What are the tipping points? What dictates the shift from normal physiological role as reservoir to ischaemic injury of splanchnic organs?
What factors exacerbate this? What can be done to prevent this?

What we do know about the splanchnic circulation

Anatomy

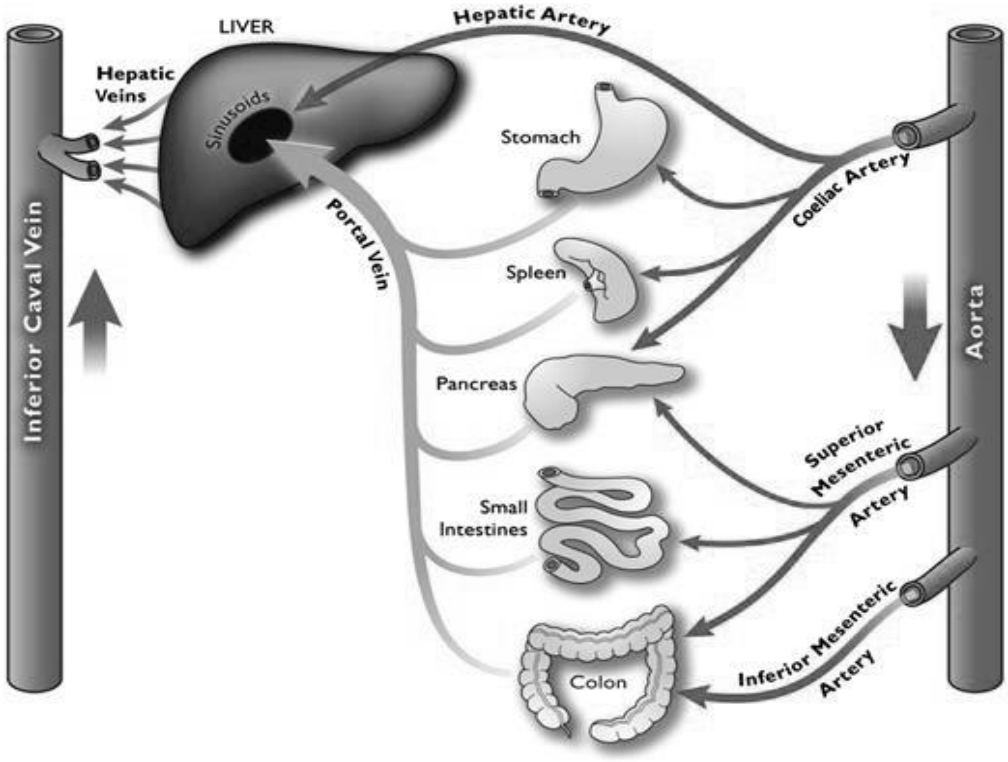


Fig. 1. Schematic representation of the splanchnic circulation.

The splanchnic system receives nearly 25% of the cardiac output through three large arteries (fig. 1): the celiac artery (usually three major branches: hepatic, splenic, and gastric) and the superior and inferior mesenteric arteries.

Main artery	Main tributary arteries	Main areas supplied
Coeliac trunk	Common hepatic (branches: right gastric and gastroduodenal)	Liver; parts of stomach, duodenum and pancreas
	Splenic	Spleen; parts of stomach and pancreas
	Left gastric	Parts of stomach and lower oesophagus
Superior mesenteric artery	Inferior pancreaticoduodenal, intestinal (12–15 branches), ileocolic, right colic, middle colic	Small intestine, caecum, ascending colon, most of transverse colon; parts of duodenum and pancreas
Inferior mesenteric artery	Left colic, sigmoid, superior rectal	Descending and sigmoid colon; parts of transverse colon and rectum

Table 1. Arterial blood supply to the splanchnic region

About 25% of the splanchnic arterial flow goes directly to the liver via the hepatic artery; the remaining 75% reaches the liver after perfusing the preportal organs. The preportal veins anastomose to form the portal vein.

Therefore the venous efflux via the portal vein represents the sum of all splanchnic arterial influx, except the hepatic arterial flow. The portal vein and hepatic artery enter the liver at its hilum and divide into progressively smaller vessels before emptying into the hepatic sinusoids. Postsinusoidal blood flows through venules, sublobular and lobular veins, and the hepatic veins, which drain into the inferior vena cava. The venous efflux via the hepatic veins represents the total hepatosplanchnic blood flow; Therefore the hepatic-arterial flow could be estimated as the difference between portal venous flow and total hepatic venous efflux.

Hepatic arterial and portal venous blood flow interact closely. Owing to this “hydrodynamic ” interaction, an alteration of flow to one of the circuits leads to an opposite change in the other circuit. The interaction tends to maintain total liver blood flow as a constant. The hepatic artery and the arteries of the preportal splanchnic organs have mean pressures of approximately 90 mmHg. The portal venous pressure is 7–10 mmHg, only slightly higher than the pressure in the sinusoids (fig 3) Most of the intrahepatic vascular resistance is distal to the sinusoids.

Splanchnic Blood Flow and Oxygen Uptake

Resting splanchnic blood flow is typically 30 ml/ min/ 100 g of tissue, which equates to 25–30% of the cardiac output. This may decrease to <10 ml/ min/ 100 g in low cardiac output states or peak locally at 250 ml /min /100 g after a meal. The splanchnic circulation must therefore be highly adaptive. Splanchnic blood flow under normal conditions is relative to the splanchnic metabolic activity and oxygen extraction capabilities of the tissues. The splanchnic oxygen consumption at rest is approximately 20–35 % of whole body oxygen consumption. When the metabolic demands of the splanchnic region increase the appropriate increase in splanchnic perfusion can be obtained by either increased cardiac output, redistribution, or by the combination of the two mechanisms. In the short term, normal metabolic processes are maintained at high levels of oxygen extraction. In extreme conditions, when hypoxia due to reduced inspired oxygen fraction is superimposed on exercise, splanchnic oxygen extraction may increase to close to 90 %. Experimental studies suggest that the gut may become hypoxic when its oxygen extraction approaches 70 % Conversely data in humans suggest that the liver is well protected against hypoxia in normal subjects, at least when hypoxic exposure is short.

	Splanchnic	Whole body	Splanchnic contribution to total (%)
Blood flow (litre min^{-1} m^{-2})	0.50–0.80	2.5–4.0	20–30
Oxygen consumption (ml min^{-1} m^{2})	20–40	110–150	20–35
Oxygen extraction fraction	0.22–0.35	0.22–0.30	

Table 2. Splanchnic and whole body blood flow and oxygen uptake at rest

The surface area of the gut mucosa is extensive for absorption (100m2) and created by villi and microvilli. It is the metabolically active area in the gut and receives over half of the total resting blood flow .The small intestinal villi are oxygenated by a countercurrent circulation where a central, small artery diffuses oxygen across to the parallel submucosal veins. At low flow rates, a substantial fraction of the O2 may be shunted from arterioles to venules near the base of the villus. The tips of the villi therefore have the lowest tissue oxygen tension (pO2), and ischaemia to this area may be induced by relatively small changes in blood flow.

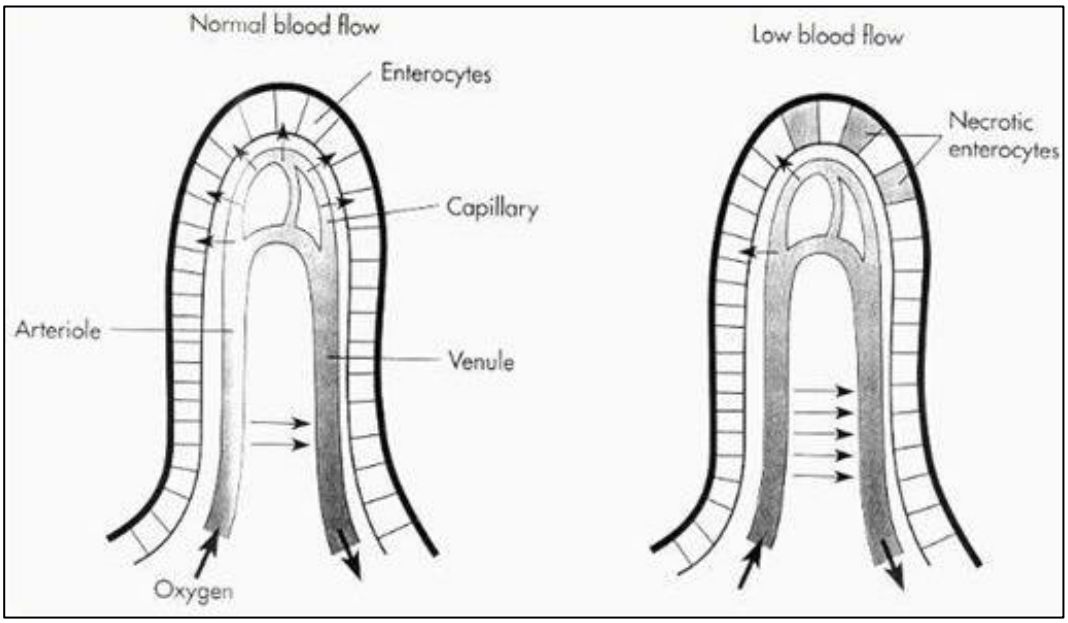


Fig 2.

The Splanchnic blood reservoir

Normovolaemic healthy adults have a blood volume of approximately 70 ml/kg body weight. Splanchnic organs constitute 10% of the body weight, but they contain 25% of the total blood volume. (+-1200 mls). Nearly two thirds of the splanchnic blood (800 ml) can be autotransfused into the systemic circulation within seconds The liver and intestines each provide between 300 and 400 ml of the blood; the spleen only contributes approximately 100 ml, but the haematocrit of this blood often approaches 75%. Therefore, the splanchnic vasculature serves as an important blood reservoir for the circulatory system. In situations of physical stress and hypovolaemia, the rich sympathetic nerve supply to the splanchnic bed causes constriction in the pre-capillary resistance vessels and post-capillary capacitance vessels resulting in prolonged reduced splanchnic perfusion. These vascular changes can occur before alterations in heart rate, blood pressure, and cardiac output. Animal studies have shown that when intestinal blood volumes range between 95% and 135% of the baseline values, cardiac output remains constant. Outside this range, constriction or dilation of intestinal vessels causes large increases or decreases in venous pressure and cardiac output to normalize these values. These observations confirm that splanchnic venous vasculature moderates changes in cardiac output during acute volume loading and haemorrhage, thereby maintaining cardiac output relatively constant over a wide range of total vascular blood volume.

Control of Splanchnic blood flow distribution

The circulation of some splanchnic organs is complicated by the existence of an **intramural circulation.** Redistribution of total blood flow between intramural vascular circuits may be as important as total blood flow.

Numerous extrinsic and intrinsic factors influence the splanchnic circulation. The intrinsic factors include local metabolic control and myogenic control, local reflexes and locally produced vasoactive substances (humoral).

The extrinsic factors include sympathetic innervation, circulating vasoactive substances (humoral) and systemic haemodynamic changes.

Alternatively one can classify the local and systemic humoral component as a third determinant. Intrinsic (local metabolic vs myogenic), extrinsic (autonomic nervous system), and humoral (local or circulating vasoactive substances). Regardless of the classification, splanchnic circulation must

always be viewed within the context of systemic haemodynamic factors. The existence of a multiplicity of regulatory mechanisms provides overlapping controls and restricts radical changes in tissue perfusion.

Intrinsic control

The autoregulatory capacity of the splanchnic bed is similar to that seen in the renal and cerebral circulations. This ensures that a constant blood flow can be maintained across a wide variety of perfusion pressures. There are two proposed mechanisms: metabolic and myogenic control.

The metabolic hypothesis focuses on oxygen supply and demand rather than blood flow. Metabolites such as H+, K+, adenosine or CO2 accumulate during periods of poor supply and tissue hypoxia and produce vasodilation restoring blood flow. Increased delivery of oxygen to the tissues will result in vasoconstriction.

The myogenic hypothesis - vessels respond to an increase in transmural pressure or stretch by constricting, restoring blood flow to baseline levels. This is mediated through opening of mechano-sensitive cation channels, mainly sodium (Na). The resulting depolarization activates voltage-gated calcium (Ca) channels inducing smooth muscle contraction. Conversely the vessels relax and reduce their tone in response to a reduction in transmural pressure.

Humoral control

Circulating vasoactive mediators of the splanchnic circulation are exogenous or endogenously produced (Table 3)

For example in postprandial hyperaemia several mechanisms are probably involved: a local reflex to the presence of luminal contents, (hyperosmolar intraluminal conditions), the release of vasoactive gastrointestinal hormones (e.g. gastrin, secretin, cholecystokinin) and the increase in gut metabolism (metabolic control through local metabolites such as adenosine and CO_2 from increased metabolic activity leading to increased blood flow).

Vasodilators	Parasympathetic tone, Increases in PCO2 , decreased PO2 and H+ ,acetylcholine, bradykinin, adenosine,gastrin,secretin,cholecystokinin,VIP,GIP,prostaglandins,substance P, leukotrienes, nitric oxide ,dopamine
Vasoconstrictors	Sympathetic tone, Decreases in PCO2 , increased PO2 and H+, vasopressin, angiotensin II, prostaglandins, peptide YY, neuropeptide YY

Table 3. Vasoactive mediators of the Splanchnic circulation

Extrinsic control

All of the splanchnic vasculature , with the exception of the capillaries receives sympathetic innervation.

The postganglionic fibres from the coeliac, superior mesenteric, and inferior mesenteric ganglia travel with the corresponding arteries. Sympathetic stimulation exerts a direct effect through the release of noradrenaline mediating α-adrenergic vasoconstriction. Alterations of blood flow in response to sympathetic stimulation follow a triphasic pattern. Initial reductions in flow return to near normal within minutes of stimulation (“autoregulatory escape”) followed by a reactive hyperaemia on ceasing of activity. Sympathetic vasoconstriction plays an important role in the distribution of blood volume throughout periods of both physiological and pathological stresses such as exercise and major haemorrhage.

Distribution and Function of Adrenergic Receptors

The hepatic artery contains α 1 α 2- β2-adrenergic receptors.
The preportal arterial supply to the splanchnic organs contains α1-, α2-, and β2-adrenergic receptors.
The preportal (intestinal) capacitance vessels have both α1- and α2-adrenergic receptors but no β2 receptors.
The portal vein contains α-adrenergic but no β2-adrenergic receptors.
Capacitance vessels of the liver (including sinusoids) have α-adrenergic receptors, and hepatic veins contain both α- and β2-adrenergic receptors.
Dopamine receptors are also present in the gut, principally DA1 and DA2, activation of which results in vasodilation

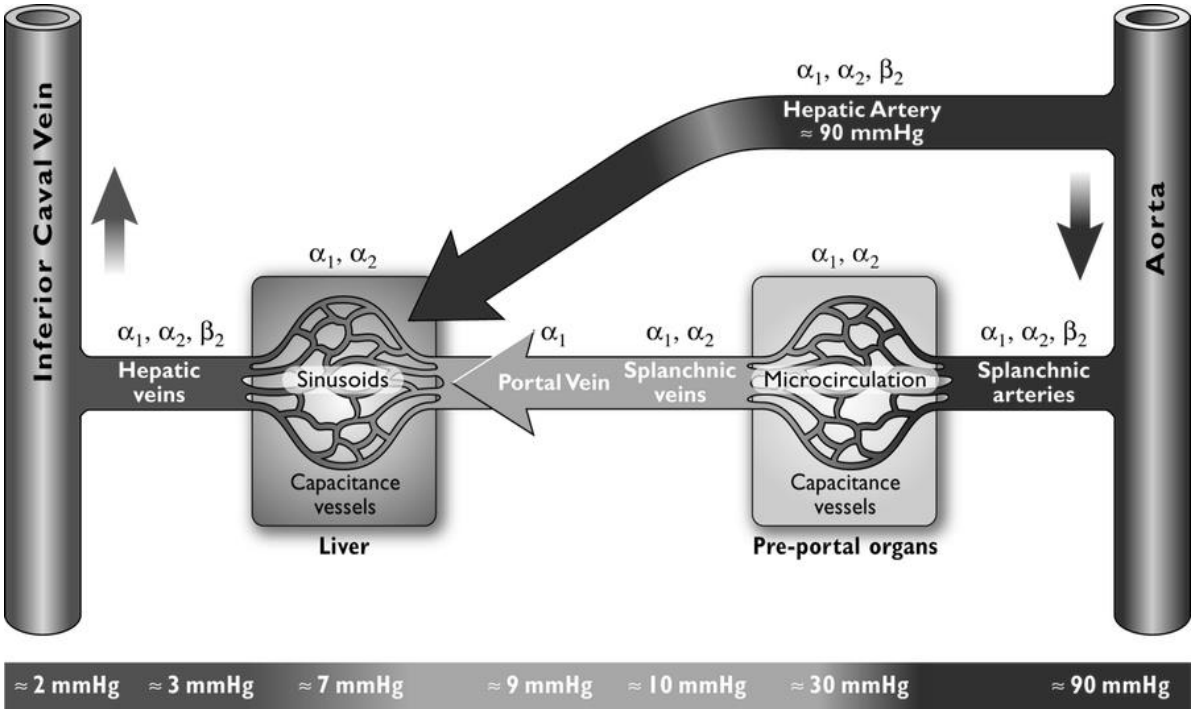


Fig 3. Diagrammatic representation of the splanchnic vasculature.

Splanchnic arteries - all arterial vessels of the preportal organs; splanchnic veins - pooled venous blood from all these organs. The distribution of adrenoceptor subtypes (α^1, α^2, β^2) and approximate intravascular pressures are shown for corresponding segments of the splanchnic vasculature.

Alpha-adrenoceptor stimulation results in vaso/venoconstriction and beta-adrenoceptor stimulation in vasodilatation. α- stimulation expels blood from splanchnic capacitance vessels, producing a rapid increase in venous return. This occurs because of active venoconstriction and passive elastic recoil of the splanchnic veins secondary to decreased arterial inflow. The initial increase in venous return may be counteracted by other α-adrenergic effects, such as an increase in hepatic venous resistance (which impedes expulsion of blood to the central circulation) and a significant decrease in splanchnic arterial inflow which results in effectively removing a portion of the vasculature from the systemic circulation by arterial vasoconstriction.

The degree to which an exogenous α-adrenergic agonist affects venous return and cardiac output is therefore dependent on many factors, including baseline myocardial contractility, blood volume, and sympathetic tone.

Pure β-adrenergic agonists augment cardiac output primarily by increasing venous return, because of increases in splanchnic blood flow due to lowered resistances in splanchnic arterial vessels and hepatic veins.

In general, a drug that stimulates both α- and β-adrenergic receptors would be expected to more effectively maintain systemic haemodynamics than one that activates either α- or β-adrenergic receptors. When simultaneously stimulated, α- and β-adrenergic receptors act together to maximally shift blood from the splanchnic vasculature into the systemic circulation by producing vasoconstriction, decreasing splanchnic vascular capacitance, and decreasing intrahepatic vascular resistance.

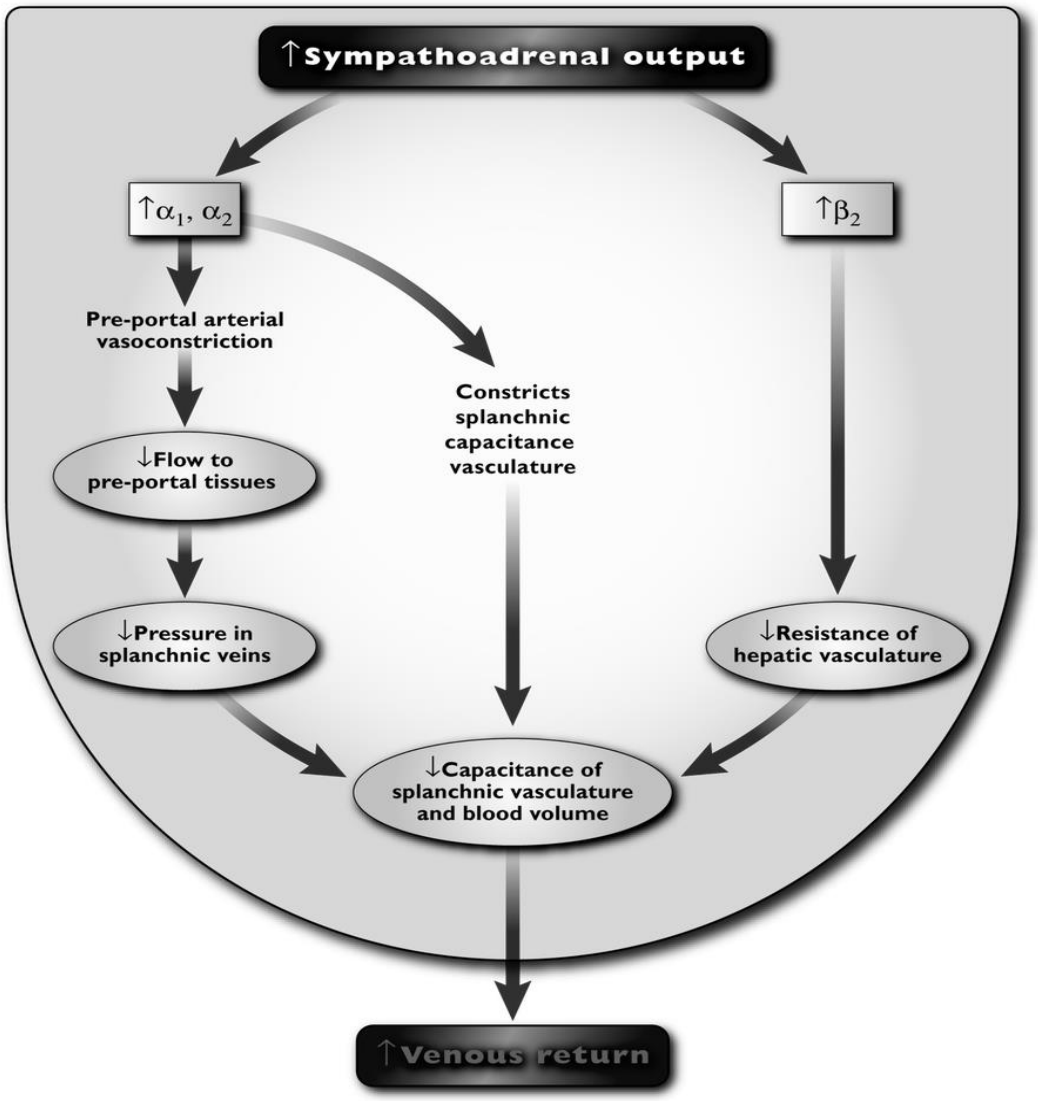
Vasoactive medications are a core component of treatment for a number of conditions where splanchnic perfusion may already be at risk. The capacity of any particular catecholamine to affect the systemic circulation through an alteration of the splanchnic circulation is determined by numerous factors, including (1) relative densities of α_1, α_2, and β_2 adrenoceptors throughout the splanchnic vasculature;(2) affinities of the catecholamine for the adrenoceptor subtypes; (3) plasma concentration of the catecholamine; (4) pre-existing tone of the splanchnic vessels; and (5) blood volume in the splanchnic vasculature

In the splanchnic system, noradrenalin has pronounced effects on both α1 and α2 receptors but has very little effect on β_2 -adrenergic receptors. Therefore noradrenalin can be expected to increase the intestinal vascular resistance, Noradrenalin is the commonest drug used for augmenting the circulation in sepsis. Its effects on splanchnic perfusion appears to be minimal. Using noradrenalin in the treatment of sepsis a further increase in mean arterial pressure (MAP) beyond 65 mm Hg does not improve or impair splanchnic perfusion. The effects of adrenaline are dose-dependent: vasodilatation at low doses and vasoconstriction at increasing doses when alpha stimulation predominates.

Phenylephrine (mainly α1) and epinephrine (α1 and β1) have been shown to reduce splanchnic blood flow, and also gastric markers of perfusion.

Dobutamine (β1 and β2), dopexamine (DA1, some β2) and low-dose dopamine (DA1 and DA2, β1 and β2, α1 in high dose) all have vasodilatory effects on the splanchnic circulation, and have been shown to improve markers of perfusion. Selective β_2 -adrenergic agonists can increase venous return by one third, entirely by increasing splanchnic blood flow and decreasing splanchnic venous resistance by more than 40%. The effect of dopamine on splanchnic circulation is complex and the information in the literature is controversial. Dopamine, at low doses, stimulates dopaminergic-1 and dopaminergic-2 receptors and produces vasodilation. At higher doses, dopamine (directly and via conversion to norepinephrine) activates both α1 and α2 adrenoceptors. Dopamine can cause hepatic arterial blood flow to decrease, increase, or remain constant. The mechanism of the decrease is unclear; it may result from the effect of dopamine on α-adrenergic receptors, particularly at relatively high doses, or from the hepatic arterial buffer response, which mediates the reciprocal relation between portal flow and hepatic arterial blood flow, when lower doses are used. The dose-related effects of dopamine on the hepatic oxygen supply-demand relation have been studied in animals. At the low end of the dose range, dopamine causes a parallel increase in hepatic oxygen supply and consumption; however, higher doses of dopamine decrease the ratio of oxygen supply to consumption. Dopamine-induced increases in mesenteric blood flow have been associated with decreases in oxygen extraction, nutritive blood flow, and capillary density in the intestines. Therefore, dopamine may divert blood flow away from splanchnic mucosa and predispose to mucosal ischemia. Low-dose infusions of dopamine were previously used as a prophylactic and therapy for acute renal failure, the DA1- and DA2-mediated vasodilation in renal and splanchnic beds was thought to be protective. This was supported by the observation that dopamine promoted a diuresis. This is now acknowledged to be a result of increased cardiac output due to dopamine's β1 and β2 effects and is no longer recommended. The addition of dobutamine rather than dopamine to a noradrenaline infusion is superior for augmenting mesenteric blood flow.

The net effects of circulating vasopressin and angiotensin are both potent intestinal vasoconstrictors. Their physiological role in the control of gut blood flow is not certain. They may both be involved in the intestinal vasoconstriction in acute hypovolaemia, and also modulate the response to sympathetic nerve stimulation and noradrenaline. Vasopressin is utilized as a second line agent in the treatment of septic shock. It exerts its effect via G-protein coupled receptors (V1). The vasoconstrictive effects of vasopressin analogues is useful in the treatment of variceal bleeding and hepatorenal failure, but it is unclear what effect this may have in the treatment of a septic patient.



Schematic representation of mechanisms by which sympathoadrenal activation can augment venous return. α^1, α^2, β^2= adrenoceptor subtypes.

Intestinal blood flow

Blood flow to the gut wall is unevenly distributed in the four main layers—the mucosa, the submucosa, the muscularis and the serosa. The mucosa and the submucosa receive up to 90 % of the flow. Changes in the total gut blood flow may influence the flow to the different layers to varying extent. The response of the intestinal vasculature to an increase in venous outflow pressure depends on the adequacy of perfusion. If perfusion is sufficient for metabolic needs, an increase in the venous pressure results in arteriolar vasoconstriction (myogenic control). If perfusion is poor and venous pressure increases, the vascular resistance decreases and the capillary density increases (metabolic control). Although metabolic and myogenic control regulate intestinal blood flow, autoregulation by increasing flow in the face of decreased arterial pressure is not as effective as that in the renal circulation. It is however enhanced in the fed state (intraluminal food present).

Parasympathetic innervations from the vagal and pelvic nerves synapse with postganglionic fibres in the gut wall and drive intestinal motility and secretions, which indirectly increase blood flow.

Hepatic blood flow

There are three principal determinants of hepatic blood flow.

1. The vascular resistance across the intestine determines portal venous flow.
2. The hepatic arterial resistance determines the hepatic arterial flow.
3. The hydrodynamic interaction or hepatic arterial buffer response regulates the interaction between the hepatic arterial and portal venous flow and is regulated by adenosine.

- The intrahepatic portal venous resistance is less important except when there is a pathological increase in portal pressure.
- Large increases in hepatic venous resistance can cause blood to pool within the splanchnic system and can lead to systemic hypovolemia.
- Autoregulation has little importance in the regulation of hepatic arterial and portal vein pressure-volume relationship. Sympathetic nervous activity is the principal neural mechanism that influences hepatic blood flow by increasing the hepatic arterial and portal resistance. The arterial vasoconstriction is transient and has similar autoregulatory escape as the intestinal vasoconstriction. In contrast, the portal response has a slower onset but the increase in resistance is sustained. In addition to the changes in vascular resistances, sympathetic nerve stimulation reduces the hepatic volume via contraction of the hepatic capacitance vessels (reservoir effect).
  Autoregulation plays an important role in distribution of blood flow within the liver (microcirculation vs macrocirculation).

Splanchnic blood flow in patients at risk of multiple organ failure
There are very few published studies with quantitative measurements of splanchnic blood flow in intensive care patients. The pathophysiology of splanchnic blood flow and inadequate splanchnic tissue perfusion in intensive care patients is multifactorial and flow abnormalities in patients at risk of multiple organ failure have not been well established. All quantitative splanchnic blood flow studies in intensive care patients at risk of multiple organ failure deal with the total hepatosplanchnic blood flow as measured by the Fick principle. Gastrointestinal tonometry cannot be used as a surrogate measure of hepatosplanchnic blood flow, since there is no consistent relationship between splanchnic blood flow and measurements obtained by gastric tonometry. Two very different patterns of changes in splanchnic blood flow and metabolic demand are common in intensive care patients:

Splanchnic Blood Flow in Low Flow States
In low flow states (e.g. cardiogenic shock) and hypovolaemia (without major injury or sepsis), splanchnic blood flow decreases without major changes in splanchnic metabolic demand. Perfusion of the heart and the central nervous system is maintained at the expense of the peripheral tissues and the splanchnic region. The effects of catecholamines administered during hypovolemia depend on the volume of blood in the splanchnic reservoir. When hypovolemia is severe, the splanchnic reservoir has already been depleted. So exogenous catecholamines ability to improve haemodynamics by modulating the splanchnic circulation decreases progressively in the face of increasing endogenous catecholamines and other vasoconstrictors (angiotensin II, vasopressin, antidiuretic hormone, endothelin-1).

Once splanchnic vasoconstriction develops in response to hypovolaemia, the recovery of the blood flow will be slow: The increased splanchnic vascular resistance and reduced blood flow persists after the circulating blood volume has been restored and systemic haemodynamics stabilized. This may increase the risk of inadequate splanchnic perfusion both after resuscitation of hypovolaemia and in intensive care patients susceptible to acute blood volume changes (e.g. as result of capillary leak). In low flow states, vasoregulation is usually well preserved and an increased oxygen extraction compensates for the reduction of blood flow. Fully developed physiological compensation may preserve adequate splanchnic tissue oxygenation even during markedly reduced hepatosplanchnic blood flow. The limits of compensation and the time of tolerance for splanchnic hypoperfusion have not been well defined. After depletion of the splanchnic reservoir, large doses of catecholamines may still increase blood pressure by further elevating arterial resistance in splanchnic and other vascular beds. Although essential for maintaining perfusion pressure and blood flow to the heart and brain the intense vasoconstriction to the splanchnic organs can produce severe ischemic injury, leading to multiorgan failure. Therefore, when treating hypovolemia-induced hypotension, the use of exogenous catecholamines should be limited to as brief a time as possible and not viewed as a substitute for the replacement of blood volume.

Ischaemia-reperfusion injury and the gut origin hypotheses
Hypoperfusion during anaesthesia and critical illness can make the gut vulnerable to non-occlusive mesenteric ischaemia. Reperfusion injury causes an increase in vascular permeability and release of inflammatory mediators into the systemic circulation and is important in the development of the systemic inflammatory response syndrome (SIRS) and multi-organ dysfunction syndrome (MODS). The majority of patients who develop MODS will at some point exhibit a septic response; but in a significant proportion no precipitating bacterial focus is found. The **gut origin hypothesis** postulates

splanchnic hypoperfusion and loss of gut barrier function occurs during the initial systemic insult. This is supported by experimental evidence showing that bacterial translocation during laparotomy or in pancreatitis, is associated with an increase in infective complications. The link between bacterial translocation and ARDS/MODS is less conclusive, with bacteria and endotoxin identified in the portal blood of only 2% of trauma patients who later went on to develop MODS. In addition, in MODS, progressive dysfunction often begins with pulmonary changes, with acute lung injury being the initial clinical picture.
The gut-lymph hypothesis may resolve this paradox. In this model, pro-inflammatory mediators from the stressed or ischaemic gut are delivered to the systemic circulation via the mesenteric lymphatics in the thoracic duct bypassing the liver to the subclavian vein and then the lungs. This theory is backed up by experimental data.

Splanchnic Blood Flow in Inflammation and Infection
In severe inflammation (e.g. systemic inflammatory response syndrome or SIRS, septic shock), the metabolic demand for oxygen in the splanchnic region is increased. In patients with normal or hyperdynamic haemodynamics, the total splanchnic blood flow is also higher than normal, but the disproportionate increase in oxygen consumption necessitates high oxygen extraction regardless of whether stable haemodynamics have been obtained by fluid resuscitation or with vasopressors.

Correction of hypotension by vasopressor drugs tends to increase the splanchnic blood flow further in hyperdynamic septic shock, endothelial injury is common in sepsis and contributes to abnormal vascular tone, blood flow, maldistribution and development of hypovolaemia. In severe sepsis systemic oxygen delivery is often further limited by acute respiratory failure. Myocardial depression is also common and can limit systemic blood flow response to the increased oxygen demand. Splanchnic hypermetabolism will increase the risk of splanchnic oxygen delivery/ demand mismatch, even with subtle changes in blood volume, cardiac output, arterial oxygenation or oxygen demand in other tissues.

During sepsis, regional hepatosplanchnic hypoperfusion may be present, even when global haemodynamic and oxygen-derived variables appear adequate. The mechanisms are controversial. **Microcirculatory circulation** is probably mediated by local metabolic and paracrine factors, and redistribution of blood between mucosal and muscularis layers has been shown. **Macrocirculatory flow** is influenced predominantly by systemic circulatory conditions such as SVR and CO. The circulatory milieu is a complex and dynamic process, and depends on stage of disease, severity, and/or therapeutic intervention.

The splanchnic circulation in liver disease
The classic disruption of the splanchnic circulation secondary to liver disease is portal hypertension (PH). PH is characterized by a pathological increase in blood pressure in the portal vein, concurrent increases in splanchnic blood flow and portosystemic collateral vessel formation. Increased blood flow in splanchnic organs augments the portal venous inflow and exacerbates portal pressure elevation leading to the formation of ascites during chronic liver disease. In addition, portosystemic collateral vessels are responsible for patients presenting acutely with a number of complications such as gastro-esophageal variceal bleeds, ascites, hepatic encephalopathy, hepatorenal syndrome, or spontaneous bacterial peritonitis, which result from blood that would normally be processed in the liver being shunted into the systemic circulation. The portal hypertensive syndrome is initiated by an increase in vascular resistance to portal blood flow at a pre-sinusoidal (portal vein thrombosis), sinusoidal (cirrhosis of the liver, chronic hepatitis, alcoholic liver disease, and hepatic schistosomiasis), or post-sinusoidal level (Budd-Chiari syndrome). The dominant cause of portal hypertension relates to liver cirrhosis, which increases resistance in hepatic sinusoids due to distortion of the liver vascular architecture with **imbalance between vasodilators and vasoconstrictors**, promoting sinusoidal constriction. A combination of increased resistance and increased flow, determines the final increase in portal pressure. The gold standard for assessing severity is the hepatic venous pressure gradient (HPVG), which is the difference between the free hepatic venous pressure and the wedged hepatic venous pressure. PH occurs when the HPVG exceeds 5 mm Hg. PH patients with liver disease often have a hyperdynamic circulation characterized by vascular vasodilatation, plasma volume expansion, and increased cardiac output which maintains and exacerbates PH. The pathophysiology of vasodilatation in the hyperdynamic circulation is multifactorial but is ultimately secondary to an elevated production of vasodilators and an abnormal response to vasoconstrictors and the formation of new blood vessels, through active neoangiogenesis under the influence of vascular endothelial growth factor.

Therapeutic Interventions

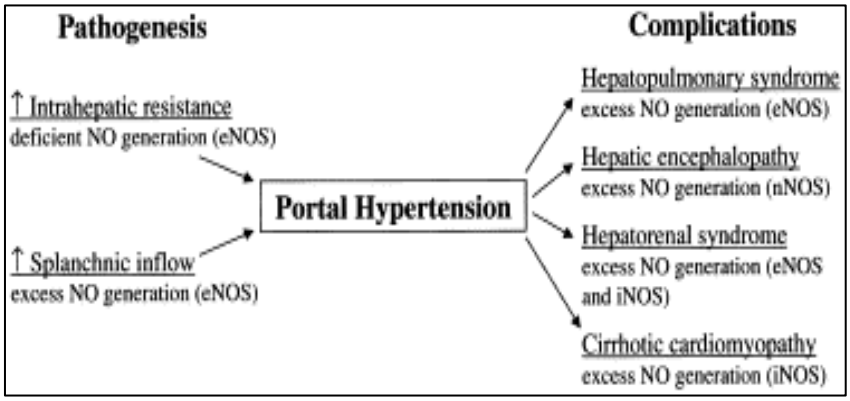
Regional anaesthesia
The relationship between regional anaesthesia, thoracic epidural anaesthesia (TEA) and SBF, has not been established. Intuitively, sympathetic block should improve flow by decreasing splanchnic vascular resistance BUT reduced SVR and cardiac output may negate this effect. In the setting of TEA, blood flow can be normalized by administration of a vasopressor but not by the administration of fluids. Colorectal surgeons, particularly, are concerned that tissue oedema due to fluid resuscitation in TEA may reduce tissue perfusion and compromise anastomoses. Evidence suggests that TEA does not increase the risk of anastomotic leak and may in fact be beneficial particularly with judicious fluid use and the use of vasopressors to manage systemic blood pressure. Human experimental data are limited, because of difficulties with direct measurement.

TEA is theoretically attractive as a therapeutic intervention for sepsis. Microcirculatory disturbance is important in the pathogenesis of sepsis, so moderating this with regional sympathetic block might influence morbidity and mortality. In experimental models of sepsis, TEA has been shown to reverse microcirculatory disturbance. The use of TEA in sepsis, either as a therapeutic intervention or to manage pain, is controversial. The concern is that the use of TEA in sepsis will increase the incidence of potentially devastating complications such as epidural abscess. However, a convincing association is yet to be demonstrated.

TEA has also been identified as having therapeutic potential in subacute pancreatitis.(SAP) Experimental models of the disease have identified improvements of pancreatic microcirculation and oxygenation that have been translated to significant survival outcomes. SAP patients are at greater risk of sepsis and coagulopathy, and may require prolonged treatment making the use of TEA challenging in this subset. Nevertheless, this technique was shown to be safe in a group of 121 patients with SAP admitted to ICU with a mortality of 2.5%.

Supplementary feeding
Enteral nutrition is an important part of the management of critically ill patients. Expert guidelines recommend that enteral feeding should be commenced early (<24 h) in all patients not expected to tolerate a full oral diet within 3 days. Under physiological conditions, oral alimentation results in a ‘postprandial hyperaemic response’, mediated primarily by local humoral and mechanical factors. The fact that this response is mirrored in critically ill patients receiving enteral nutrition raises concerns in the haemodynamically unstable patient. The first being the potential for triggering bowel ischaemia by an increase in oxygen demand exceeding supply, the second being the potential for the steal phenomenon whereby gut blood flow increases at the expense of other ‘vital’ organs. However early enteral nutrition appears beneficial even in patients requiring vasopressor support. In contrast, SBF appears to be reduced during parenteral nutrition.

Nitric Oxide
Experimental studies show that nitric oxide is important in the maintenance of basal vasodilatation in the mesenteric vasculature and the hepatic artery. Nitric oxide is synthesized from L-arginine by nitric oxide synthase (NOS) which exists in several isoforms including eNOS (endothelial nitric oxide synthase).Nitric oxide activates guanyl cyclase to create cyclic GMP which then mediates vascular smooth muscle relaxation via protein kinase G. The role of NO in splanchnic blood flow is evidenced by the demonstration that inhibitors of NOS decrease splanchnic blood flow by inhibiting NOS at the mesenteric arteriolar level.



In addition to increasing splanchnic inflow, eNOS-derived NO generated from liver endothelial cells increases flow within the sinusoidal channels of the liver itself. NO has also been implicated in hepatic microcirculatory and splanchnic vascular dysfunction, particularly in portal hypertension. Although mechanical factors contribute majorly to increased liver resistance in portal hypertension, there is a vasculogenic component to the development and perpetuation of this syndrome as well. The vascular component of portal hypertension includes an increase in splanchnic blood flow, and an increase in intrahepatic vascular resistance. Dysregulation of the NO system plays a role in both these processes. The increase in intrahepatic resistance is mediated through a deficiency in eNOS-derived NO generation in the liver. Excess eNOS-derived NO generation at the mesenteric arteriolar level increases splanchnic blood flow. Excess generation of NO has also been implicated in hepatopulmonary syndrome, hepatic encephalopathy, hepatorenal syndrome, and cirrhotic cardiomyopathy.
Therapeutic approaches aimed at manipulating the NO pathway in portal hypertension and its components have been an area of investigation in both experimental systems and in humans and have met with varying success.

Ventilation and IPPV
Intermittent Positive pressure ventilation (IPPV) may impact on cardiac output mainly by reducing preload, thereby reducing splanchnic perfusion. This is exacerbated by the use of very large tidal volumes, high levels of positive end expiratory pressure, and high inspiratory pressures. This is probably due to increased hepatic venous pressures and mesenteric vascular resistance, with reduced portal blood flow. At normal ventilator pressures, adverse effects are minimal. Recruitment manoeuvres worsen splanchnic oxygen delivery, despite improving arterial oxygenation. Spontaneous breathing efforts during airway pressure release ventilation improve cardiac output and splanchnic perfusion. Prone ventilation does not affect splanchnic perfusion, if intra-abdominal hypertension is avoided. Permissive hypercapnia helps to maintain splanchnic perfusion.

Measurement of Splanchnic Blood Flow
Human studies on splanchnic blood flow present methodological difficulties attaining quantitative measurements. Direct measurement of splanchnic blood flow and intravascular pressures are practically difficult due to anatomical reasons (multivessel influx and efflux, mixing of portal venous and hepatic arterial blood within the liver).
Therefore much of the basic splanchnic circulatory physiology has been extrapolated from experimental studies and has not been confirmed in normal humans. Total hepatosplanchnic blood flow can be estimated according to the **Fick principle** from the hepatic uptake of substances that are exclusively metabolized by the liver and distributed in the plasma. Hepatic venous catheterization is necessary for the measurement.

Splanchnic perfusion can be measured in a variety of ways; however, because of a variety of issues, monitoring has not become common in clinical practice.

1. The use of indicator dilution with Indocyanine green (ICG) was first described more than 40 yrs ago. This original application of this technique required hepatic venous catheterization to measure the post-hepatic concentration of ICG. From this the application of the Stewart–Hamilton equation can be used to calculate SBF; flow equals the amount of injectate divided by the area under the post-hepatic concentration curve. ICG is eliminated from the blood solely by hepatocytes. This process can be monitored noninvasively using the principle of pulse

spectrophotometry and the optical absorption of ICG with a transcutaneous probe, allowing for assessment of liver function and splanchnic perfusion.

2. Gastric tonometry has been extensively studied as a monitor of splanchnic perfusion. This technique needs a gastric catheter to measure the concentration of carbon dioxide in the stomach and arterial measurements of carbon dioxide and bicarbonate. Using the Henderson–Hasselbalch equation the gastric intramucosal pH (pHi) can be calculated.

$$\mathrm{pH_i} = 6.10 + \log\left(\frac{\mathrm{HCO_3\ arterial}}{\alpha \cdot \mathrm{pCO_2 gastric}}\right)$$

where pHi is the gastric mucosal pH and α is the solubility of CO_2

Alternatively the concentration difference between gastric and arterial carbon dioxide concentration (PgCO2) can be calculated. Both pHi and PgCO2 are prognostic indicators in the critically ill but do not correlate well with splanchnic perfusion.

3. Although not directly part of the splanchnic circulation, assessments of the sublingual circulation have now been developed. This utilizes its greater accessibility within the gastrointestinal tract. The use of sublingual capnometry allows sublingual CO_2 to be measured ($PslCO_2$) which has been correlated with lactate concentrations and outcomes. At present there is no clear evidence to support its widespread use.

4. Laser Doppler flowmetry utilizes the frequency shift (Doppler principle) in a laser beam to measure flow. This has been used in trials to establish colonic blood flow, as have ultrasound-based Doppler measurements.

Evaluation of the splanchnic microcirculation is complicated by the inability to directly measure blood flow and oxygenation in human disease.

The monitoring of splanchnic oximetry through the use of near-infrared spectroscopy (NIRS) is an emerging method used to assess tissue oxygenation status. Although splanchnic tissue oxygenation $SrSO_2$ is thought to be potentially of high value in critically ill patients, $SrSO_2$ has been found to have a relatively high degree of variability that can potentially make it difficult to interpret. In addition, currently available sensors only penetrate a few centimetres deep. Because splanchnic organs only lie near the skin surface in children and infants it can be difficult to use clinically in a non-invasive manner in adults. Splanchnic oximetry does hold great promise in the ability to monitor patient oxygenation status and detect disease states in humans, especially in paediatric populations. Researchers examining its clinical use in adults have bypassed the problem of abdominal wall thickness by placing miniature NIRS sensors directly on the organs of interest, rather than on the skin surface above.

Conclusion

Much of what we know about the physiology of the splanchnic circulation has been extrapolated from experimental studies because routine assessment of SBF in humans is impractical. Techniques for direct measurement are invasive while reliability is an issue with indirect alternatives. As such a thorough understanding of the splanchnic circulation during the perioperative period and critical illness remains elusive. Despite these problems, manipulation of splanchnic physiology is a fascinating area of research and therapeutic techniques for targeting the splanchnic circulation, such as RA in sepsis or SAP need further investigation.

References

1. Takala J. Determinants of splanchnic blood flow. British Journal of Anaesthesia (1996); 77:50-58
2. Sean M, Bailey J. Review of splanchnic oximetry in clinical medicine. Biomed Opt (2016); 21:091306
3. Gelman S, Phillip S, Mushlin M. Catecholamine-induced Changes in the Splanchnic Circulation Affecting Systemic Hemodynamics. Anesthesiology 2004; 100:434-9
4. Harper D, Chandler B. Splanchnic circulation. BJA Educ 2015; 16:66-71.
5. Fernandez M, et al. Pathogenesis of Portal Hypertension: Extrahepatic Mechanisms. Curr Hepatology Rep 2016; 15:199–207

Pulse Oximetry
Principles and practice

Prof. Jenny Thomas

UCT Dept of Anaesthesia & Perioperative Medicine
Emeritus Professor

INTRODUCTION

The unreliability of cyanosis in the recognition of arterial hypoxaemia was described in the classic paper in 1947 by Comroe and Botelho. Pulse oximetry's value lies in its non-invasive ability to differentiate between pulsatile arterial blood flow and the smooth-flowing capillary and venous flow. This makes pulse oximetry (PO) the ultimate safe monitor in detecting hypoxaemia, and thus is now mandatory for anaesthesia provision in most countries in the world. Its use in other clinical settings such as emergency medicine and intensive care is ever-increasing. Recent studies in its use prior to discharge of neonates from maternity units, has directly impacted on the surveillance for congenital heart disease.

The term "pulse oximetry" refers to the way that the measurement is picked up (using the pulse wave) and to the technology of oximetry required to provide the readings. Pulse oximetry is a combination of two technologies: photospectrometry and optical plethysmography. **Photospectrometry** is the technique whereby the saturation of haemoglobin is estimated by the computer of a pulse oximeter. **Optical plethysmography** focuses the measurement on pulsatile arterial flow.

HISTORY

The **history** of oxygen saturation monitoring goes back to the mid-1850s. The relationship between the absorption of light and the amount of absorbance was first described by Johan Heinrich Lambert (1728–77) and published in 1760. Nearly 100 years later, these principles were further investigated by August Beer and published as the Beer-Lambert Law in 1851. The following is a list of important developments leading to where oximetry is today.

1864: Georg G Stokes discovers a pigment that is the oxygen carrier in blood.
1864: Felix Hoppe-Seyler purifies the pigment and calls it haemoglobin.
1876: Karl von Veirordt studies the reflection spectra of haemoglobin solutions and the finger
1887-90: Carl Gustav Hufner studies absorption spectra
1919: August Krough and I Leicht use spectroscopic methods to measure oxygen saturation of blood of fish
1931: Ludwig Nicolai investigates the quantitative spectrophotometry of light transmitted through human tissues
1934: Kurt Kramer makes precise measurements of the oxygen saturationof blood flowing through cuvettes.
1935: David Drabkin and James Harold Austin measure the spectrum of undiluted haemolysed and non-haemolysed blood
1939-45: World War II: great military interest in oximetry of pilots at high altitude
1940: JR Squires passes red and infrared light through the finger web for the continuous monitoring of oxygenation. This required continuous compression of tissues to create a bloodless field for calibration.
1940-42: Glen Allan Millikan coins the term oximeter and develops the Millikan oximeter.
1948-50: Earl Wood develops Wood's ear oximeter
1960: Development of the first bench "CO-oximeter", able to distinguish between haemoglobin, carboxyhaemoglobin, and methaemoglobin.
1964: Robert Shaw develops the first eight wavelength ear oximeter
1970: Hewlett-Packard market the eight wavelength ear oximeter
1971: Takuo Aoyagi uses the pulsatility of the absorption signal to separate absorption due to the arteries from other tissues
1974: Aoyagi develops the prototype pulse oximeter using an incandescent light source, filters and analogue electronics
1975: First commercially available pulse oximeter.

The Hewlett-Packard ear oximeter was developed for respiratory physiology studies, not for clinical use. It used eight different wavelengths from an incandescent light source with narrow band interference filters, and transmitted to the ear pinna and thence via fibre optics to a detector. No separation of absorption from arteries, veins or capillaries was made. Despite being a huge advance in oximetry monitoring, it had a large

ear piece, a stiff connection between ear and main monitor, and needed frequent calibration. It was not user-friendly and has now been replaced by the pulse oximeter discussed here.

COMPONENTS OF A PULSE OXIMETER

Probe: Energy light source (LED)
Photo detector: picks up the amount of absorbed light
Computer and cables: (MRI-compatible PO has optical fibre transmission of signals)
Display screen: plethysmogram, heart rate, alarms, volume control.

OPTICAL PRINCIPLES: Conventional pulse oximeters.

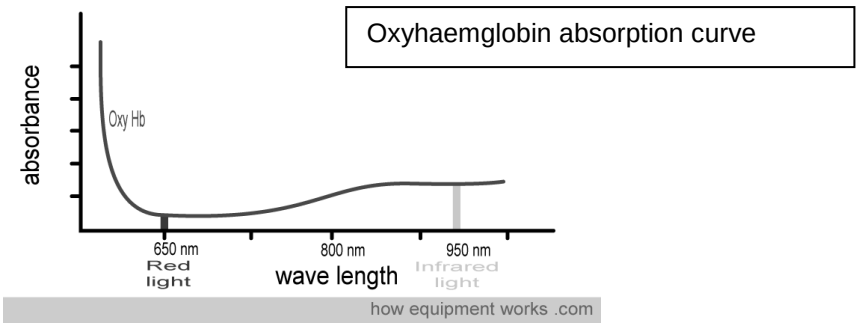
Haemoglobin: To say that haemoglobin consists of four polypeptide chains belies the highly complex three-dimensional shape of the molecule. The two α chains and two β chains have 141 and 146 amino acid residues respectively. Each polypeptide chain is combined with one haem group and each haem group has one atom of iron in the ferrous state. This iron atom is able to combine reversibly with one molecule of oxygen and yet always remain in the ferrous state. The haem groups are attached to the chains through histidine by weak bonds and are positioned in crevices in the chains. The quarternary structure of the 4 chains seems to form a crumpled ball but the actual shape is critically important to pulse oximetry. This shape not only controls the reaction with oxygen, but also, any change in the shape of the quaternary structure with the level of oxygenation, alters the optical absorption spectrum proportionately. This is the basis of pulse oximetry.

Spectrophotometry (aka photospectrometry): The concentration of any transparent substance in solution may be measured spectrophotometrically. In PO, light is emitted from a light source (usually a light emitting diode), goes across the PO probe and reaches a light detector. If you place a finger between the light source and detector, light will have to pass through this finger to reach the detector. Part of the light will be absorbed, and that not absorbed will reach the detector. The amount of light absorbed by the finger depends on many properties that the PO uses to calculate oxygen saturation. Three important physical properties are:

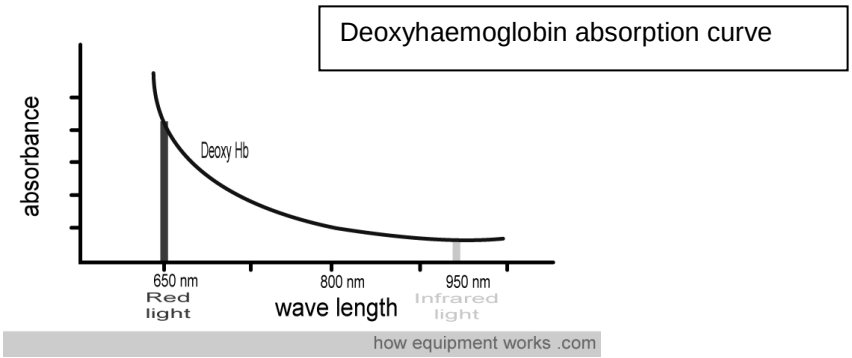
- *The concentration of the light absorbing substance*
  The amount of light absorbed is proportional to the concentration of the light absorbing substance (in this case haemoglobin). This is **Beer's Law**. i.e the more haemoglobin, the more light is absorbed.
  By measuring how much light reaches the detector, the PO knows how much light has been absorbed. The more haemoglobin in the finger, the more the light is absorbed.
- *The length of the light path in the absorbing substance*
  The amount of light absorbed is proportional to the length of the light path in the absorbing substance (haemoglobin concentration=amount of Hb per unit area). This is **Lambert's Law**. If one compares a wide artery vs a narrow artery, although the Hb per unit area is the same, the wider artery will provide more Hb particles in the way of the light source than the narrow one (less path length less absorption, more path length more absorption).
- *The fact that oxyhaemoglobin and deoxyhaemoglobin absorb red and infrared light differently.*
  Oxyhaemoglobin absorbs more infrared light than red light. Deoxyhaemoglobin absorbs more red light than infrared light. Why is this important?

All light is composed of waves, where a wave length is the distance between the tips (or troughs) of the wave. Light wave lengths vary with different colours and are specific for their colour, but are very short, with the unit of measurement being a nanometer (1m =1,000,000,000nm). 650 nm is a shorter wavelength than 940nm. PO uses the property that oxyhaemoglobin and deoxyhaemoglobin absorb light of different wave lengths in a specific way.

Absorption spectra for different types of haemoglobin
For PO, a light source sequentially passes light of different wave lengths through a sample of oxyhaemoglobin. The detector notes how much light has been absorbed at each wave length. From this, a graph of the absorbance of **oxyhaemoglobin** at different wave lengths shows it does not absorb the same amount at each different wave length, with more light being absorbed at the infrared band than at the red band.

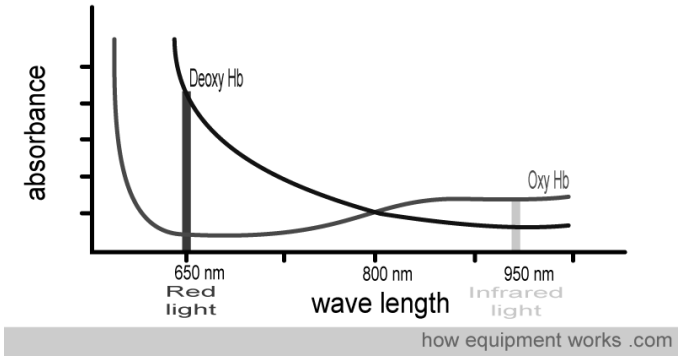


The same can be shown for deoxyhaemoglobin, and its absorption is different from that of oxyhaemoglobin



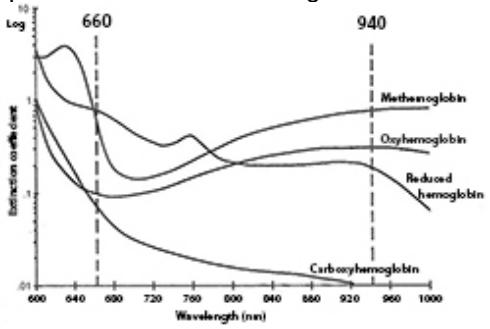
Reduced Hb (deoxyHb) absorbs more light in the red band (and less in the infrared band)

Oximetry thus has two lights to analyse: red and infrared. From the above two diagrams, it can be seen that oxyHb absorbs more red light than deoxyHb, and deoxyHb absorbs more infrared than oxyhaemoglobin. (Although depicted in red and blue, infrared is invisible to the human eye and is not blue).



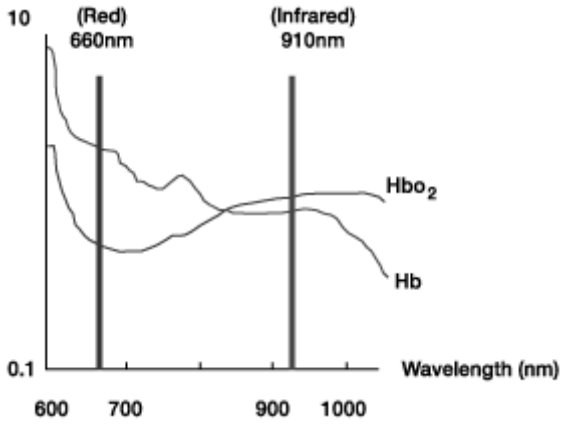
Pulse oximetry calculates the oxygen saturations by comparing how much red light and how much infrared light is absorbed by the blood. Depending on the relative amounts of oxyHb and deoxyHb present at the time of the reading, the ratio of the amount of red light absorbed compared with infrared light absorbed changes. Using this ratio, the PO is able to work out the oxygen saturation.
At 0% saturations, there is only deoxyHb present. At 75% saturations, both dexoyHb and oxyHb are present and at 100%, only oxyHb is detected.

Comparisons of the absorption spectra of different haemoglobins



Isobestic point: The wavelength at which the absorption spectra of two species cross each other i.e **th**e point at which two substances absorb a certain wavelength of light to the same extent. It is a specific wavelength, wavenumber or frequency at which the total absorbance of a sample **does** not change during a chemical reaction or a physical change of the sample. The word derives from two Greek words: "iso", meaning "equal", and "sbestos", meaning "extinguishable".

In oximetry, the isobestic points of oxyhaemoglobin and deoxyhaemoglobin occur at 590 nm (not seen on this curve) and 805 nm. These points may be used as reference points where light absorption is independent of the degree of saturation. Some earlier oximeters corrected for haemoglobin concentration using the wavelength at the isobestic points.



Thus comparison of absorbencies at different wavelengths allows estimation of the relative concentrations of HbO and Hb (i.e. saturation). Modern pulse oximeters may use two or more wavelengths, not necessarily including an isobestic point.

CALIBRATION ADJUSTMENTS (graphs or curves)

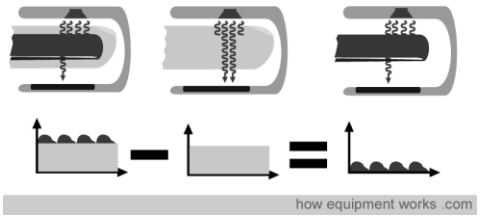
Oximeters are calibrated during manufacture and automatically check their internal circuits when they are turned on. They are accurate in the range of oxygen saturations of 70% to 100% (+/-2%), but less accurate under 70%. Due to the shape of the oxyhaemoglobin curve, the saturation starts to fall rapidly at 90%. Beer and Lambert's Laws require that light passing through a substance should go straight through (a single absorbent, no reaction between the absorbent and the solvent, and no possibility of a photochemical reaction). This is not possible in blood because of substances eg red blood cells, causing scatter. This necessitates a calibration graph to correct for errors. This was originally done where volunteers were exposed to varying inspired oxygen concentrations and their blood samples were taken against this, but where their saturations were lower that 75%, it was felt that this was no longer safe and thus unethical. Below 75%, calibration curves are based on mathematical estimations, thus readings are typically less accurate below these levels. Each pulse oximetry producer has a calibration module specific to their oximeter. More recent POs have aimed to provide more accurate readings for lower saturation clinical settings such as cyanotic congenital heart disease. Calibration is based on CO-oximetry which is a bench type oximeter.

The Beer-Lambert Law: relates the fraction of radiant energy absorbed by the substance to the concentration and amount of the substance

$$\mathbf{A + \log (I_0/I) = \varepsilon L c}$$

A = absorbance of sample, and is proportional to the concentration (c) and the path length (L) or depth of the sample.
ε is the molar extinction coefficient which is a wavelength dependent constant characterising the sample. (It is defined as the optical density of an absorbing substance in a concentration of 1 mmol/l measures with a light path length of 1 cm at a specific wave length). I_0 is the intensity of the energy without the sample, and I is the intensity with the sample.

Plethysmography (aka optical plethysmography)
In a digit or an ear (common body parts for measuring oxygen saturations), arterial blood is not the only thing that absorbs light. Skin and other tissues e.g. fat, do too. It is necessary to be able to ignore these in the calculations of oxygen saturations. Pulsatile arterial flow provides a constantly changing absorbance against the non-changing absorbance (non-pulsatile) of other tissues. The computer is able to mathematically extract "changing absorption" from the total signal, leaving only that image on display i.e pulsatile arterial blood flow. This pulsatile signal is a very small percentage of the total signal (about 2%) and it is because of this small amount of the total image, that PO is very sensitive to errors due to movement, poor probe placement, and ambient light interference. A poor trace will fool the computer into making a faulty reading.



Plethysmography trace: PO shows this pulsatile change in absorbance as a graph. This image provides a considerable amount of information and this tracing should be evaluated before looking at the number provided as a percentage saturation.

AC and DC components refer to these pulsatile and non-pulsatile components.

- *Non-pulsatile component* is the non-pulsatile blood and tissue pigmentation which produces a direct current (DC).
- *Pulsatile component* is provided by the pulsation of the artery which produces an alternating current (AC).
- Light absorption during diastole is only from the deoxygenated components (tissues etc).
- Light absorption during systole provides an increase in both wavelengths, and these pulse-added absorbances are thus caused by the Hb in arterial blood.

The SpO_2 value: Algorithm
The **fractional** oxygen saturation (% HbO_2) is the ratio of oxyHb to the sum of all the Hb species present, whether available for reversible binding to oxygen or not.
Functional oxygen saturations (SpO_2) is defined as the ratio of oxyHb to all functional haemoglobins.

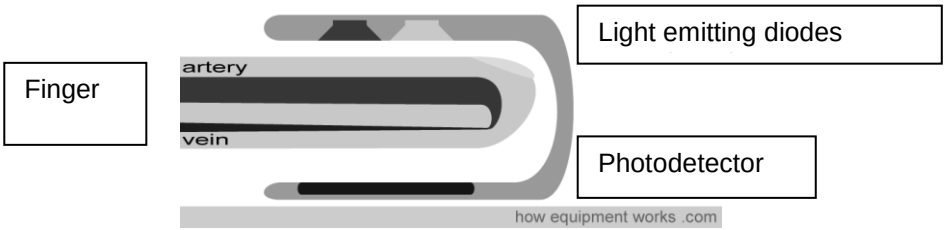
Functional vs Fractional
HHb =0 HbO_2 = 96% COHb = 2% MetHb = 2%
Functional saturations (SaO_2) = 96 / 96+ 0 = 100%
Fractional saturations (HbO_2) = 96/ 96+2+2 = 96%
Most manufacturers choose to use a functional saturation in creating an algorithm
Absorption ratio (R):

$$R = \frac{AC_{660} / DC_{660}}{AC_{940} / DC_{940}}$$

HOW PULSE OXIMETRY WORKS
Oximetry by any **spectrophotometric** method relies on a change of colour or more accurately, a change in the absorption of electromagnetic energy, with a change in the percentage of oxygen bound to the haemoglobin molecule.
The functioning of conventional oximeters compares the absorption of energy at two wavelengths, usually 660 and 950 nm, passed through an extremity. A value, SpO_2, which is approximately equal to arterial haemoglobin saturation, SaO_2, is determined from the ratio of the absorption of the energy at the two wavelengths. The range of wavelengths that may be used in vivo are from about 600 to 1300nm; wavelengths shorter than 600, red skin pigment (melanin) causes a high level of absorption whereas at greater than 1300nm, there is strong absorption due to water in the tissues.

The PO performs spectrophotometry either by **reflection** (from the skin and subcutaneous tissues) or **transmission** (through an extremity). Most PO in clinical practice are transmission PO. In the design of the PO, wavelengths must be chosen that are easy to generate, ideally monochromatic, and low in cost. A sufficiently sensitive detector is required so that high energy levels are not needed, and then some computing power is needed to extract a saturation values from all the other absorbants. Light emitting

diodes (although not strictly monochromatic) are now used as the energy source and have the advantage that they can be switched on and off rapidly, so it is possible to use a single detector if the diodes are alternatively switched on and off.



Light source: Light emitting diodes
Advantage of LEDs over other energy sources

- Narrow band width (almost monochromatic)
- High efficiency
- Low temperature
- High switching speeds (>1MHz)
- Stable peak wave length
- Intensity varies linearly with drive current

There are two LEDs at the light source side of the probe. Both are never lit at the same time.
The two LEDs cycle ON and OFF between 2000-3000 times/second, and the PO switches them on and off in a particular sequence:
Red light activated → through finger to detector → stray light also to detector → detector records red light plus room light → red light off. Then…
Infrared light on → through finger to detector → room light detected by detector → detector reads both → Both LEDs switched off.
Only light read = room light recorded by detector
Computer subtracts room light from reading and finds red and infrared light levels, calculating the SpO_2.

Electronics of a pulse oximeter: summary

- Amplification of the photodetector signal
- Separation of the red and infrared plethysmography signal
- Switching and controlling the current of the light emitting diodes
- Adjustment of the gain of one of the two signals to make them equivalent
- Separation of the "arterial" (pulsatile) composition of the signal
- Analogue to digital conversion of the red and infrared signals
- Calculation of the red : infrared ratio
- "Calculation" of the oxygen saturation (SpO_2)
- Display: SpO2, plethysmogram , heart rate
- Control of alarms
- Storage of trends

LIMITATIONS
The oximeter averages its readings every 10-20 seconds. Hence, they cannot detect acute desaturation.
The finger probe has a response time of approximately 60 seconds, whereas the ear probe has a response time of 10-15 seconds.
The site of application should be checked at regular intervals, as pressure sores and burns have been reported.
The pulse oximeter only provides information about oxygenation. It does not give any indication of the patient's carbon dioxide elimination.

Safety	
Technical	*Physiological*
Mechanical / motion artefacts	Pulse dependency
Electromagnetic interference	Pulse volume
MRI	Pulse rhythm
Dangerous	
Technical	*Physiological*
Accuracy	Abnormal Hb
Calibration	Other absorbents and pigments
Flooding	Dyes
Penumbral	Delay
	Pulsatile veins

Artefacts:

- Toxic alteration in haemoglobin structure
- CO Hb (carboxyhaemoglobin): Reduction in assessment of SpO_2 by pulse oximeter (i.e. overestimates the fraction of Hb available for oxygen transport)
- CyanmetHb: No information
- Methaemoglobin: At high levels, SpO_2 reads at about 85%, independent of actual Hb.
- SulfHb: Not reported (but affects Co oximetry by producing falsely high levels of MetHb.)

Structural haemoglobin

- HbF: No significant effect
- HbH: No significant effect (overestimates the fraction of Hb available for O_2 transport)
- Hb Kohn: Artefactual reduction in SpO_2 of 8 to 10%
- HbS: No significant effect

Dyes

- Fluorescein: No significant effect
- Indogo carmine: Transient decrease
- Indocyanine green: Transient decrease
- Isosulfan blue: Low dose no effect; high dose produces a prolonged reduction in SpO2
- Methylene blue: Transient, marked decrease in SpO2 lasting up to several minutes; possible secondary effect as a result of haemodynamic consequences

Haemoglobin substitutes: Diaspirin cross-linked Hb or Bovine polymerised Hb: No effect.

Hb Concentration:

- Anaemia: No effect with normal SaO_2. If hypoxia with Hb < 14.5g/dl, progressive underestimation of actual SaO_2.
- Polycythaemia: No effect

Put differently:

No effect:

Foetal haemoglobin (HbF), SulphHb, Bilirubin (absorption peaks are 460, 560 and 600 nm), dark skin.
Jaundice: No effect (multi-wave laboratory oximeters may read falsely low saturations and falsely high COHb and MetHb)
Acrylic fingernails: No significant effect.

Falsely low reading:

Methaemoglobin (MetHb). The presence of MetHb will prevent the oximeter from working accurately and the readings will tend towards 85%, regardless of the true saturation.
Methylene blue. When methylene blue is used in surgery (e.g. parathyroidectomy or to treat methaemoglobinaemia), a short-lived reduction in saturation estimations is seen. Readings may fall by 65% at a concentration of 2-5 mg/kg for between 10 and 60 minutes.
Indocyanine green. Use of this dye (e.g. in cardiac output studies) may cause a transient reduction in recorded saturations.
A reduction in peripheral pulsatile blood flow produced by *peripheral vasoconstriction* results in an inadequate signal for analysis.

Sensor contact: "optical shunting" of light from source to detector directly, or by reflection from skin causes a falsely low SpO2.
Venous congestion, which may be caused by pulsatile tricuspid regurgitation, high airway pressures and the Valsalva manoeuvre, may produce venous pulsations which can produce low readings.
Venous congestion of the limb may affect readings, as can a badly positioned probe.
External fluorescent light in the operating theatre may cause the oximeter to be inaccurate, and the signal may be interrupted by surgical diathermy. Shivering may cause difficulties in picking up an adequate signal.
Nail varnish, especially blue varnish, may cause falsely low readings.

Falsely high reading

Carboxyhaemoglobin (CoHb). CoHb (haemoglobin combined with carbon monoxide) is registered as 90% oxygenated haemoglobin and 10% desaturated haemoglobin - therefore the oximeter will overestimate the saturation.
Ambient light: bright or flickering light (close to the harmonic of the LED) may falsely increase SpO_2.

SUMMARY and conclusion

Pulse oximetry:

- Based on the physical principles of the Beer- Lambert Law and different light absorption spectra of oxyHb and deoxyHb.
- Estimates SpO_2 from the differential absorption of red (660nm) and infrared (940nm) light in the tissues
- These two wavelengths allow differentiation of reducedHb (deoxyHb) and oxyHb: reduced Hb absorbs more light in the red band than oxyHb, and oxyHb absorbs more light in the infrared band
- The computer of the PO computes the ratio between the two signals and relates this ratio to the arterial saturation, using an empirical algorithm
- Spectrophotometry may be by reflection from skin and subcutaneous tissues or by transmission through an extremity.
- Is essential for safe perioperative care.

References and recommended reading

1. Comroe JH, Botelho S. The unreliability of cyanosis in the recognition of arterial hypoxaemia. Am J Med Sci 1947; 214:1-6.
2. Howequipmentworks.com. 2017.
3. Moyle JTB, Hahn CEW, Adams AP. Pulse oximetry: Principles and Practice Series. BMJ 1994.
4. *Anaesthesia UK: Principles of pulse oximetry:* www.frca.co.uk/article.aspx?articleid=332
5. WHO working manual on Pulse Oximetry:
6. http://www.who.int/patientsafety/safesurgery/pulse_oximetry/en/index.htmland at
7. http://www.who.int/patientsafety/safesurgery/en/index.html
8. Fearnley SJ. Pulse Oximetry. Update in Anaesthesia :http://www.nda.ox.ac.uk/wfsa/html/u05/u05_003.htm
9. Hill E, Stoneham MD. Practical applications of pulse oximetry. http://www.nda.ox.ac.uk/wfsa/html/u11/u1104_01.htm
10. Principles of pulse oximetry. http://www.oximeter.org/pulseox/principles.htm

Near-Infrared Spectroscopy

Dr Lance Thompson

Private Practice & Honorary Lecturer
UCT Dept of Anaesthesia & Perioperative Medicine

INTRODUCTION

(Please note: most of the content will focus on cerebral oximetry, with only short explanations of application to other organ sites)

Maintaining adequate oxygen delivery to tissues and organs (especially the brain) is a primary goal of anaesthesia and critical care. Hypoxia and reduced oxygen delivery are deleterious, but the brain is still one of the least monitored organs.

Cerebral oximeters are non-invasive, continuous monitoring devices, used to monitor adequate cerebral oxygenation. They utilise similar physical principles to pulse oximeters. In 1977, Jobsis first reported the use of near-infrared spectroscopy (NIRS) to measure cerebral oxygenation. The use of cerebral oximetry has the potential to improve outcome in certain surgical population groups. In addition, cerebral oxygenation desaturation leads to adverse peri-operative outcomes.

THE MONITOR AND EQUIPMENT[1]

Only the INVOS (IN Vivo Optical Spectroscopy) will be considered here (Covidien, Dublin, Ireland, formerly Somanetics Corporation, Troy, Michigan, USA), as this is the most common monitor in the country. This measures the regional cerebral oxygenation ($rScO_2$). The NIRO-200 (Hammamatsu Photonics, Hammamatsu City, Japan) is another monitor; it measures the tissue oxygen index (TOI)[5].

The cerebral monitor consists of a monitor connected to oximeter probes. Adhesive pads attach probes to the patient's scalp overlying the frontal lobe. Probes comprise a fibreoptic light source and light detectors. Light sources release light in the infrared range through light emitting diodes. This emitted light reaches the underlying cerebral tissue. Reflected light (influenced by the oxygenation status of haemoglobin) returns to the surface and is detected by the light detectors in the oximeter probes. A $rSCO_2$ value will then be displayed on the monitor screen.

2-Channel Monitor


Array of INVOS™ Sensors

NIRS PRINCIPLES[2,3]

1. Light in the NIR spectrum (700 – 1300nm) penetrates biological tissue several centimetres. This is due to the relative transparency of tissue to light in this wavelength range. With cerebral oximetry, NIR light can penetrate the skull to underlying cerebral tissue (superficial cortex).

2. Certain biological molecules (chromophores – light-absorbing molecules) have distinct absorption spectra in the NIR.
    a. The primary chromophores are metal complexes: haemoglobin, bilirubin and cytochrome (also melanin)
    b. Absorption spectrum of deoxygenated haemoglobin (Hb or HHb)): 650 – 1000 nm
    c. Absorption spectrum of oxyhaemoglobin (HbO_2): 700 – 1150 nm
    d. Commercial devices utilise wavelengths between 700 – 850 nm where the absorption spectra of Hb and HbO_2 are maximally separated and there is minimal overlap with water (which would start absorbing photons above 950 nm)
    e. The isobestic point (wavelength at which 2 molecules have the same molar absorptivity) for Hb/ HbO2 is 850 nm – this can be utilised to measure total tissue haemoglobin concentration

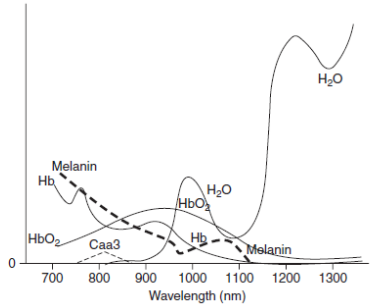


Fig 1 Absorption spectra for oxygenated haemoglobin (HbO_2), deoxygenated haemoglobin (Hb), Caa3, melanin, and water (H_2O) over wavelengths in NIR range. Note the relatively low peak for Caa3.

(JM Murkin and M Arango, BJA 2009; 103 (PGA Supplement): i3 – 9.)

Beer-Lambert Law(1,2,3)

Cerebral oximeters calculate cerebral oxygenation using the Beer-Lambert equation, which is a combination of two physical laws

Beer's Law
The intensity of transmitted light decreases exponentially as the concentration of a substance the light passes through increases.
Thus, as the concentration of a substance increases, the amount of light absorbed by the substance increases and the amount of light detected by the photodetector decreases.

Lambert's Law
The intensity of transmitted light decreases exponentially as the distance travelled by the light through a substance increases.
Thus, as the distance a light travels through a substance increases, the amount of light absorbed increases, and the amount of light detected by the photodetector decreases.
An amount of a substance can thus be determined by how much light the substance absorbs.

The Beer-Lambert Law may be expressed as:

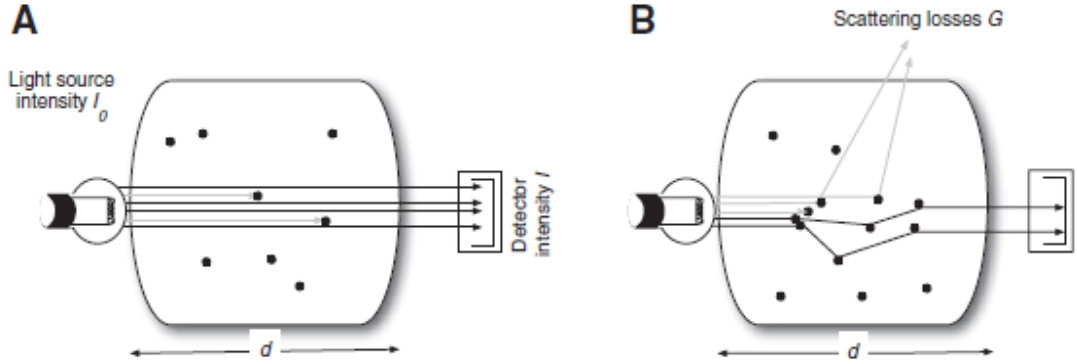
$$\Delta A = L \times \mu$$

ΔA - amount of light attenuation
L - differential photon pathlength through tissue
μ - absorption co-efficient of chromophore X, expressed as [X] x ε
[X] - tissue concentration of chromphore X
ε - extinction co-efficient of chromophore X
[X] = ΔA/ L x ε (may allow measurement of SO_2)

Alternatively, it may be expressed as:

$$\mathbf{A = I/ I_0 = \varepsilon.c.d}$$

A – light attenuation
c – chromophore concentration
d – source- detector distance
ε - absorption co-efficient (thus attenuation is also directly proportional to this)
I – incident light intensity
I_0 – detected light intensity



(A Gosh et al, Anesthesia Analgesia 2012; 116 (6): 1373-1383.)

In biological tissue light scatter also becomes a major contributor to attenuation by two mechanisms:

1. Scattering losses (not all emitted light reaches the detector)
2. Multiple scatter of some of the light, resulting in a greater distance travelled than the source-detector distance (see figure B)

Differential spectroscopy, a modification of the Beer-Lambert law, aims to address the effects of scattering:

$$\mathbf{A = I/IO = \varepsilon.c.d.DPF + G}$$

DPF - differential path length factor (must be defined first – difficulties in determining)
G – scattering losses (assumes that these losses remain constant)

NIRS TECHNIQUES[3, 4]

These rely on a measure of optical attenuation (total loss of light caused by absorption and scattering). There are different approaches to derive the resultant physiologically relevant signals:

1. Differential spectroscopy
2. Multidistance spectroscopy (or Spatially resolved spectroscopy) – used in the INVOS
    - An array of closely spaced detectors is used to measure light attenuation as a function of source-detector separation
    - Combining these measures with an estimation of the wavelength dependency of light scattering, a scaled absolute haemoglobin concentration can be derived, i.e. the relative proportions of HbO_2 and HHb
    - By measuring only the ratio, the need to estimate the optical path-length is avoided
    - An oxyhaemoglobin saturation can then be calculated:
    - **$RSO_2 = HbO_2/ (HbO_2 + HHb)$**
3. Frequency resolved (or domain) spectroscopy
4. Time-resolved spectroscopy

ALGORITHMS[3]

Cerebral oximeters use mathematical algorithms to translate measured changes in light attenuation to a physiological measure (e.g. changes in HbO2, HHb and tissue oxygen saturation). The INVOS series apply propriety algorithms in order to do this. Further algorithms involving subtraction of values from the emitters near and far from the photodetector are used to limit contamination from extracranial blood. There are other commercially available cerebral oximeters. Inter-device variability occurs due to different wavelengths of light emitted by the probes, different light sources and different mathematical algorithms.

NIRS Limitations and confounding factors[2]

1. Extracerebral Tissue

- Transcutaneous NIRS is reflective of a heterogeneous field comprising vascular and non-vascular tissue
- For cerebral oximetry, photons must penetrate several tissue layers (scalp, skull and dura)
- Increasing transmitter/ receptor distance increases the depth of penetration, but power must be limited to prevent thermal damage

2. Spatial Resolution

- Mean depth of photon penetration is approximately 1/3 the transmitter/ receiver separation
- By using two differentially spaced receiving optodes, a degree of spatial resolution can be achieved
- The proximal or shallow receiver detects mainly superficial tissue, whereas the distal or deep detector reflects deeper tissue
- Subtracting the proximal from the distal value cerebral RSO_2 at a depth of 1-2 cm is obtained
- About 85% of cerebral rSO_2 is derived from cortical tissue and 15% from overlying extracerebral tissue

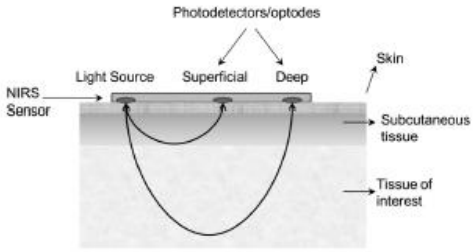


Fig. 2. IN Vivo Optical Spectroscopy (INVOS) System (Covidien).

3. Cerebral Arterial/ Venous Blood Partitioning

- NIR cerebral oximetry does not rely on pulsatile flow (unlike pulse oximetry)
- Cerebral oximetry measures a weighted average of arterial, capillary and venous compartments
- For cerebral cortex, average tissue haemoglobin is about 70:30 for venous: arterial blood volume (capillary volume of 2% is ignored)
- There may be considerable variation in individual cerebral arterial/ venous ratios
- Significant haemodilution may also affect readings
- Thus the use of cerebral oximetry as a trend monitor may minimise confounding effects (with interventions aiming to keep values close to individual baseline values)

4. Extracerebral Haematoma

- Extradural or subdural haematomas can change the proportion of cerebral to extracerebral haemoglobin, and thus affect values
- There is also the potential for artefact and signal attenuation when extracerebral tissue is thickened or oedematous

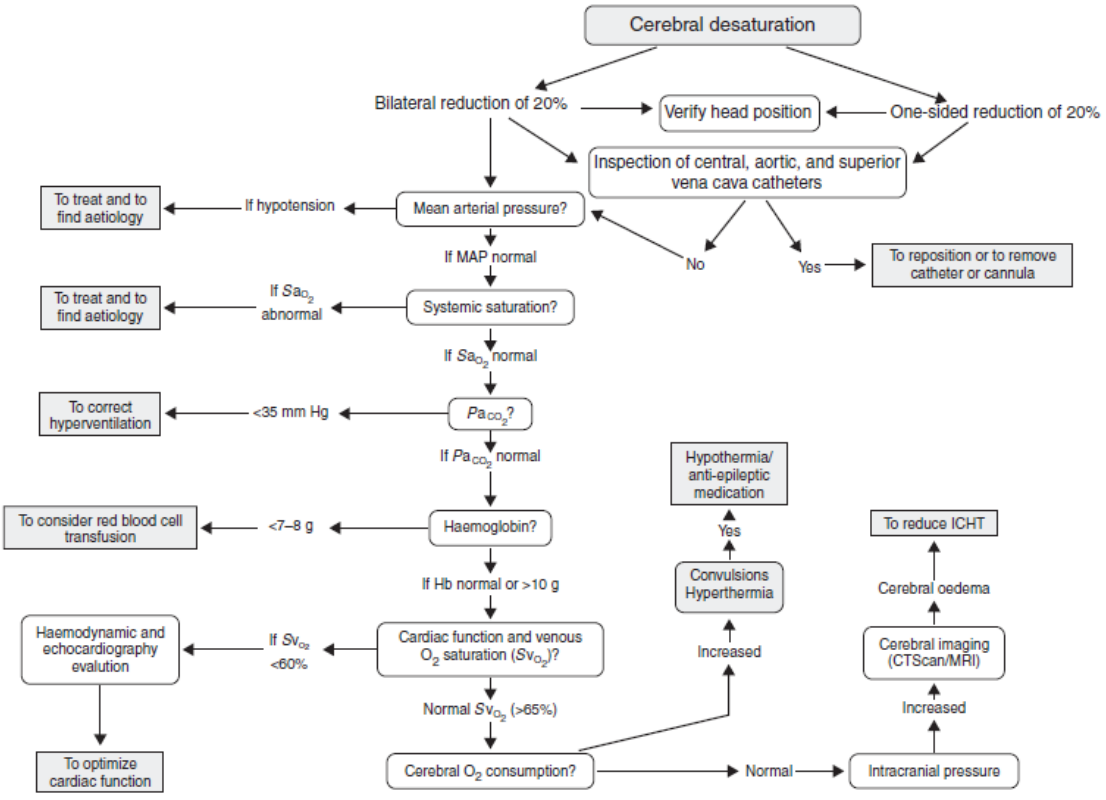
5. Non-haeme Tissue Chromophores

- Melanin pigmentation (found in hair) can significantly attenuate light transmission and thus impede NIR measurements
- Thus probes are placed on the frontal eminences away from the hair line
- Melanin in skin does not appear to interfere significantly
- Conjugated bilirubin (absorption peak of 730 nm) may attenuate the NIR signal
- Even at high levels changes in cerebral perfusion can be detected (supports the use of establishing baseline values)

NORMAL VALUES

- There is wide intra- and individual baseline variability in rSO_2
- Normal range: 60 – 75% (co-efficient of variation 10%)
- Thus best used as a trend monitor
- Changes in rSO2 are referred to as either:
  - Percentage point reduction, or
  - Percentage reduction from baseline
- No validated cerebral rSO_2 derived ischaemic thresholds
- Initiating triggers in clinical studies and management protocols:
  - Absolute cerebral rSO2 value: </ = 50% of baseline
  - >/ = 20 % reduction from baseline

CLINICAL INTERPRETATION OF CEREBRAL OXIMETRY (1)



Denault et al. Semin Cardiothoracic Vasc Anaesth 2007; 11: 274-81.

Adequate cerebral oxygenation is dependent upon adequate cerebral blood flow and oxygen content.

FACTORS RESULTING IN REDUCED CEREBRAL rSO_2	
CEREBRAL BLOOD FLOW	**OXYGEN CONTENT**
Cardiac output	Hb concentration
Acid-base status	Hb saturation
Major haemorrhage	Pulmonary function
Arterial inflow/ venous obstruction	Inspired oxygen concentration

A common limitation in cerebral oximetry monitoring has been the lack of intervention protocol to treat decreases in rSO_2. Denault et al devised one such algorithm, included above.

GOAL-DIRECTED THERAPY AND NIRS (6)

Cerebral oximetry may give a target for monitoring oxygen delivery (DO_2) to the brain (goal-directed therapy). For most organs, oxygen delivery is a function of cardiac output and arterial oxygen saturation. The uneven distribution of systemic vasoconstriction makes the brain somewhat protected against decreases in cardiac output. Thus cerebral blood blow becomes a function of cerebral perfusion pressure (CPP), not cardiac output:

$$\mathbf{DO_2\,(brain)\ \alpha\ (CPP\ x\ r4\ x\ [Hb]\ x\ \%Sat)\ /\ \eta}$$

r – arterial resistance vessel radius
η - blood viscosity

The following table depicts an alternative way how aspects of oxygen delivery and consumption can be viewed to prompt interventions when cerebral oximetry is low:

Table 1 Goal-directed therapy guided by NIRS

Variable	Clinical scenario causing cerebral desaturation	Intervention
O_2 delivery		
CPP	Hypotension, elevated ICP, elevated central venous pressure	Maintain ABP above the lower limit of autoregulation. Check venous drainage cannula
r^4	Hypocarbia, vasospasm, malpositioned arterial cannulae	Decrease minute ventilation, pH-stat management. Check aortic cannula position
[Hb]	Anemia	Transfusion
% Sat	Cyanosis	Lung recruitment maneuvers, increase F_iO_2, manage Q_P/Q_S ratio
η	Polycythemia, sickle-cell disease	Partial exchange transfusion, permissive anemia
O_2 consumption	Fever, seizure, arousal	Cooling, sedation

CPP, cerebral perfusion pressure; ICP, intracranial pressure; ABP, arterial blood pressure; *r*, resistance vessel radius; [Hb], blood concentration of hemoglobin; % Sat, arterial oxygen saturation; η, blood viscosity.

Kasman N, Brady K. Ped Anesth 2011 (21).473 – 478.

FURTHER CLINICAL APPLICATIONS (5,7)

Global or regional changes in DO_2 create a state of oxygen debt and anaerobic metabolism with failure to meet metabolic demand (oxygen consumption, VO_2). The severity of the oxygen debt may be linked the development of morbidity and mortality. Prompt interventions directed at reversing oxygen debt as associated with improved outcomes. Unfortunately, standard vital signs paramaters are not highly predictive of circulatory failure. Furthermore, biochemical markers of organ hypoperfusion (blood lactate, unmeasured anions, base deficit) may correlate with the severity of oxygen debt, but cannot be measured continuously and lag circulatory changes.

Systemic venous oxygen saturation (SvO_2) monitoring provides an estimate of global oxygen balance (using the Fick equation, where $\mathbf{SvO_2 = S_aO2 – (VO_2 / DO_2)}$). This measurement is invasive, however.

With NIRS, Regional oxygen saturation (rSO_2) approximates regional SvO2, and in combination with arterial oxygen saturation allows for estimation of regional oxygenation. The regional Fick equation (**rSO2 = SaO2 – (VO2/ DO2**) may be manipulated to derive:

- Regional arterio-venous difference: **ΔarSO2 = SaO2 – rSO2**
- Fractional oxygen extraction (fOE): **fOE = [SaO2 – rSO2]/ SaO2**
- An **increase in FOE** may be due to:
    - Reduced delivery with constant brain oxygen consumption
    - Higher oxygen consumption than delivery
- A **decrease in FOE** may be due to:
    - Decrease in oxygen extraction due to decreased utilisation
    - Constant oxygen consumption with an increased oxygen delivery

NIRS can thus provide regional circulation monitoring which may drive organ specific goal-directed therapies.

Further clinical limitations of cerebral oximetry[1]

1. Blood from an extracranial source
2. Electrosurgical equipment (diathermy) can affect accuracy
3. Only regional cerebral oxygenation is measured. Large areas of the brain remain unmonitored
4. Unable to identify a cause for the desaturation

Clinical applications[1]

(Not an exhaustive list)

1. Cardiac surgery
   a. CABG
   b. Deep hypothermic circulatory arrest

2. Vascular surgery
   a. Carotid endarterectomy
   b. Carotid endarterectomy hyperperfusion syndrome

3. Orthopaedic surgery
   a. Shoulder surgery (beach chair position)

4. Paediatrics
   a. Cardiac surgery
   b. Periventricular leucomalacia
   c. NICU – premature infants (PDA, ventilation)

5. Additional Uses
   a. Extra cerebral sites
   b. Hepatic
   c. Renal
   d. Splanchnic
   e. Somatic (muscle bed)
   f. Free flap surgery

These additional sites show promise especially in the paediatric population, where major organs are close to the skin

CONCLUSION

Cerebral oximetry is a simple, real-time, continuous, non-invasive monitoring system that may improve patient outcomes in different clinical settings. As always, future research is required to validate the use of cerebral oximetry monitoring (and oximetry at other sites) in bringing about enhancements in patient outcomes.

References

1. Tosh W, Patteril M. Cerebral Oximetry. BJA Education 2016, 16(12):417-421.
2. Murkin JM, Arango M. Near-infrared spectroscopy as an index of brain and tissue oxygenation. Br J Anaesth 2009; 103 (Suppl.):i3-13.
3. Gosh A, Elwell C, Smith M. Cerebral Near-Infrared Spectroscopy in Adults: Work in Progress. Anesth Analg 2012; 115:1373-83.
4. Sood BG, McLaughlin K, Cortez J. Near-infrared spectroscopy: Applications in neonates. Seminars in Fetal and Neonatal Medicine 2015 (20): 164-172.
5. Toet MC, Lemmers PMA. Brain monitoring in neonates. Early Hum Dev 2009; 85:77-84.
6. Kasman N, Brady K. Cererbral oximetry for pediatric anesthesia: why do intelligent clinicians disagree? Ped Anesth 2011 (21):473-478.
7. Scott JP, Hoffman GM. Ped Anesth 2014 (24): 74-88.

Peripheral Nerve Stimulators
Physical Principles and Clinical Applications

Dr Aneet Kessow

Private Practice & Honorary Lecturer
UCT Dept of Anaesthesia & Perioperative Medicine

Peripheral Nerve Stimulators (PNS) are used:

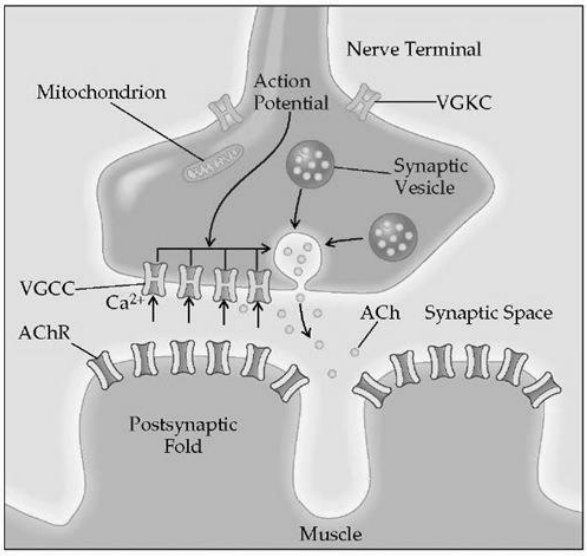
To monitor neuromuscular blockade

- During induction of anaesthesia for intubation
- During surgery to guide repeated doses of muscle relaxants to assess the depth of relaxation
- To differentiate between different types of block
- At the end of surgery to assess the ability of reversal
- At the end of anaesthesia or in recovery to assess the degree of residual blockade

For peripheral nerve identification

- To help identify peripheral nerves for nerve mapping
- To help identify peripheral nerves for nerve blockade
- To help prevent intraneural nerve injection of local anaesthetic

An understanding of neuromuscular junction electrophysiology is vital to understanding the clinical application of PNS.



Why is monitoring so important?

Undetected Residual neuromuscular block is common in the recovery room. Up to 42 % of patients receiving intermediate acting non-depolarizing muscle relaxants arrive in with a TOF ratio of less than 0.7. Standard clinical criteria do not reliably detect residual blockade and standard of care has shifted the accepted TOF ratio to 0.9 or greater. Residual neuromuscular blockade contributes to greater adverse respiratory events, greater chance of re-intubation, longer recovery stay and ICU admission.

Properties of an electrical nerve stimulator

The PNS generates a standard electrical pulse, which should be *Supramaximal* – A supramaximal stimulus is a stimulus that is greater than that needed to activate all the nerve fibres in a nerve. A stimulus of 20-25% more than that needed to generate a maximal clinical response is usually used. Electric stimulation is the most commonly used but theoretically magnetic can also be used.

Constant current not voltage generator
A current setting of 60 MA will achieve supramaximal stimulation in most cases. The level of current passing through the nerve between electrodes, not the voltage, achieves nerve stimulation. The resistance in the patient's skin is influenced by:

- Electrode gel dries or makes poor contact
- The patient's level of expiration changes
- Changes in patient body temperature – resistance increases as the skin temp cools

Current delivered by constant current stimulators remains the same even when the patient's resistance changes compared to constant voltage stimulators, thus ensuring supramaximal stimulation. Current output should be limited to less than 80 MA to prevent tissue damage.

The *Rheobase* is the minimal current required to stimulate the nerve with a long pulse.

Monophasic rectangular waveform
Monophasic and rectangular waveforms are essential to prevent multiple nerve stimulations. Assuming a square pulse of the current is used to stimulate the nerve, the total energy (charge) applied to the nerve is a product of the intensity of the current and the duration of the pulse.

Duration of stimulus

Chronaxy is the property whereby a current must flow for a minimum time before the nerve tissue will depolarize, no matter how high the current. The Chronaxy time is about 80 µsec in mammalian motor nerve.

The Chronaxy can be used as a measure of the threshold for any particular nerve and it is useful when comparing different nerves or nerve fiber types. Certain nerves have a different chronaxy based on their physical properties (myelination, size, etc). Also, certain patient conditions, such as diabetes, have an effect on chronaxy.

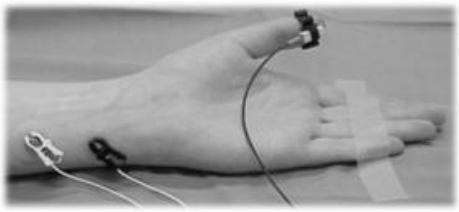
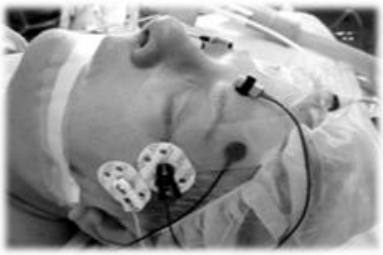
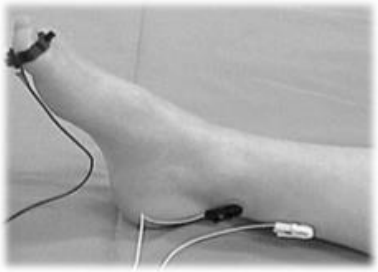
Safety and ergonomic features

- Portability
- Warning when the set current is not delivered
- Display of polarity – so that the negative electrode can be placed distally for maximal response.
- Different stimulus patterns available
- Default to lower current options when attached to needle electrodes or a stimulating needle

Placement of Electrodes:

The electrodes should be placed over the path of the peripheral nerve. The negative (black) electrode is the activating stimulating electrode, and the most effective stimulation is obtained when placed closest to the muscle terminus. The positive electrode (red) is placed 2cm proximally.

Common sites for placement appear below:

DIAGRAM	NOTES
	Nerve: Ulnar nerve Muscle: Adductor pollicis Action: Thumb adduction Black: 1-2cm proximal to wrist crease Red: 2-3cm proximal to black
	Nerve: Facial nerve Muscle: Orbicularis oculi and Corrugator supercilii Action: Twitching of eyelid and eyebrow Black: Just anterior to tragus Red: Lateral to outer canthus of eye
	Nerve: Posterior tibial nerve (sural nerve) Muscle: Flexor hallicus brevis Action: Plantar flexion of great toe Black: Over posterior aspect of medial malleolus, over posterior tibial artery Red: 2-3cm proximal to black

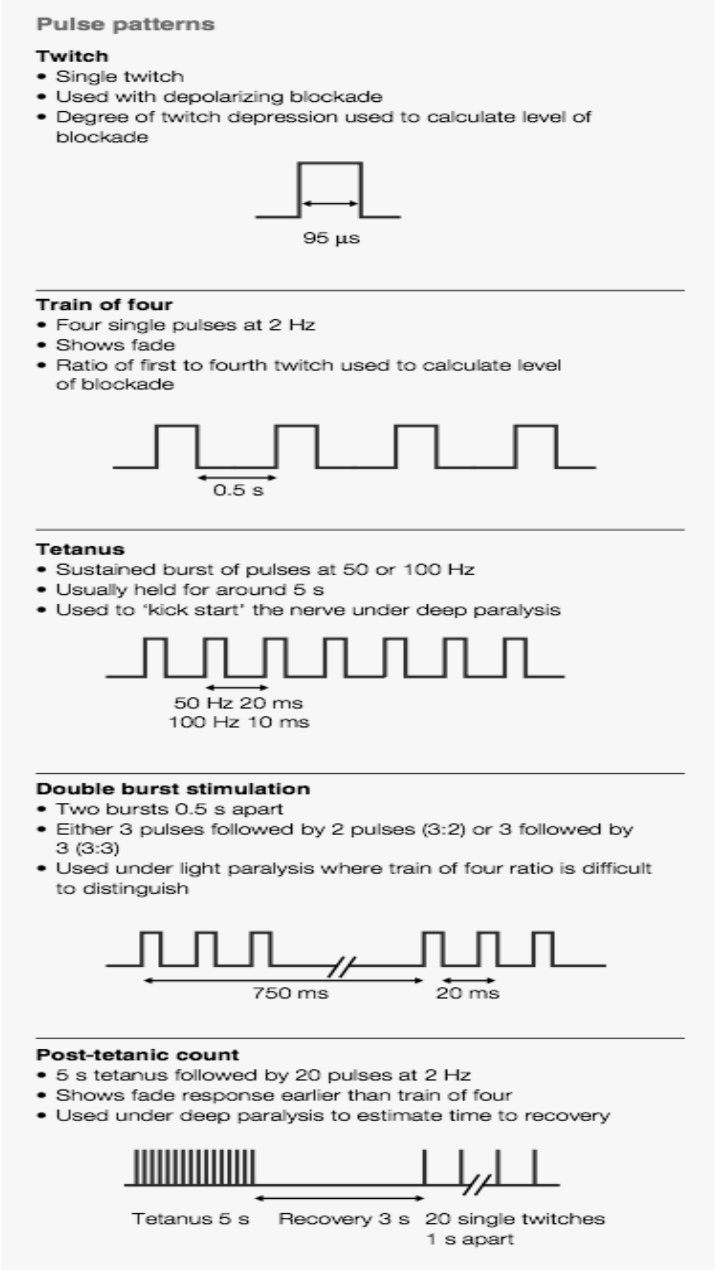
2 sites are commonly used to establish the depth of paralysis. Measurement at the adductor pollicis correlates well with the tone in the upper airway and upper oesophageal muscles. The muscles around the eye (orbicularis oculi and corrugator supercilii) recover early (similar to the diaphragm) and a TOF of 4 here is often correlated with a TOFC of 2 or less at the adductor pollicis.

Patterns of stimulation

Single twitch
An electrical pulse is delivered at 1Hz, and the ratio of the evoked twitch compared with that before muscle relaxation gives an indication of neuromuscular blockade. When 75% of the post junctional Ach receptors are occupied by NMBA, twitch magnitude starts to decrease. When there is 100 % occupation, no twitch is elicited.

Useful for monitoring of post junctional receptor, as for deep relaxation required for intubation.

Train of four
Four pulses are given at a frequency of 2Hz or 0.5 seconds apart, potentially eliciting 4 twitches (T1-T4). The ratio of the first to the last (T1:T4) indicates the degree of neuromuscular block. Non depolarizing neuromuscular blocking agents (NMBA) occupy receptors producing a decrease in magnitude of the first twitch compared with a pre relaxant stimulus and a progressive decrease in magnitude of T1 compared to T4 (Fade). At 75% occupancy T4 disappears. Similarly at 80, 90 and 100 % occupancy T3, T2 and T1 disappear. With recovery the twitches reappear in reverse order.



TOF ratio is the ratio of T1 vs. T4
TOF count is a numerical count of twitches elicited with a TOF stimulus

Clinically accepted values for TOF count:

1 twitch for intubation
1-2 twitches during anaestheisa3-4 twitches before reversal of neuromuscular blockade

Double Burst Stimulation (DBS)
This consists of 2 bursts of 3 stimuli at 50 Hz with each triple burst separated by 750 ms. These bursts appear visually as 2 separate stimuli T1 and T2, and the ratio of these is related to the TOF ratio. It is used as it is easier for the operator to interpret reliably compared to comparing the T1:T4 ratio visually.

Tetanic stimulation
The impulse at 50 Hz of 5ms produces detectable Fade in muscle contraction, the extent of which is related to the neuromuscular block. No fade indicates no neuromuscular block. In intense neuromuscular block, TOF stimulation may elicit no twitches. Tetanic stimulation causes post-tetanic facilitation (PTP) to mobilize presynaptic Ach. Subsequent 1Hz twitches can now overcome the high concentration of NMBA. The number of twitches generated (ie the post tetanic count) reflects the degree of neuromuscular blockade.

Depolarizing NMBAs react differently to the PNS modes of stimulation. They produce equal but reduced twitches in response to single twitch and TOF stimulation (the T4:T1 ratio is 1), and reduced but sustained contraction with tetanic stimulation. They do not demonstrate either tetanic fade or PTP.

How do we monitor the effect of the stimulus?

The muscle response can be assesses by visual and tactile methods – these are the easiest, but can often be unreliable and inaccurate. Objective measurement of neuromuscular monitoring is the only way of accurately assessing residual neuromuscular blockade. It is conducted via quantitative measurement of the strength of contraction of a peripheral muscle in response to peripheral nerve stimulation produced by 2 stimulating electrodes. Each measurement technique measures the force of contraction either directly or by a factor proportional to that force.

Electromyography uses electrodes to record the evoked electrical response of the muscle. Stimulating electrodes are placed over the nerve and recording electrodes over the muscle being stimulated. Stimulation of the nerve results in depolarization of the muscle and the amplitude of the compound muscle action potential is recorded and expressed as a percentage of control or as a TOF

ratio. Typically the ulnar nerve is used and the electrodes are placed over the muscle of adductor pollicis. This is easily accessible but a drawback is that small movements of the hand may affect the response, as may electro cautery.

Mechanomyography uses electrodes to measure the evoked muscle response of a muscle. A small weight is suspended from the muscle to maintain isometric contraction. The tension produced on PNS is converted into an electrical signal. This gold standard is very accurate, but is difficult to set up in practice and is used mainly in research.

Acceleromyography uses Newton's second law of motion. Force equals mass times acceleration. The transducer uses a piezoelectric crystal secured to the distal part of the digit (thumb) being measured and the PNS provides the electrical stimulus. (ulnar nerve). Acceleration of the distal digit is directly proportional to the force of the contraction (mass stays the same) and therefore inversely proportional to the degree of NMB.

Kinemyography uses a piezoelectric polymer sensor in the groove between the thumb and the index finger. Movement of the muscle generates a voltage in the sensor which can be measured. This is not commonly used.

Phonomyography uses a high fidelity narrow bandwidth microphone placed alongside the muscle, and measure sound intensity. This is not commonly used.

Clinical evaluation of responses after administration of Depolarising NMBA

Phase 1 Block

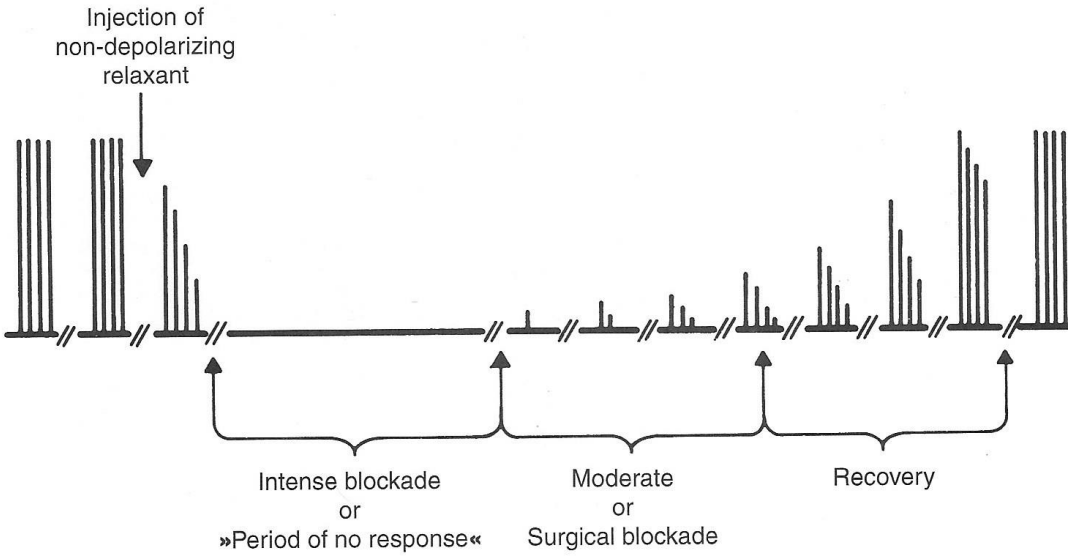
- Patients with normal plasma cholinesterase activity who receive a moderate dose of succinylcholione (0.5-1.5mg/kg).
- There is no fade during tetanic stimulation or TOF
- No PTP

Phase 2 Block

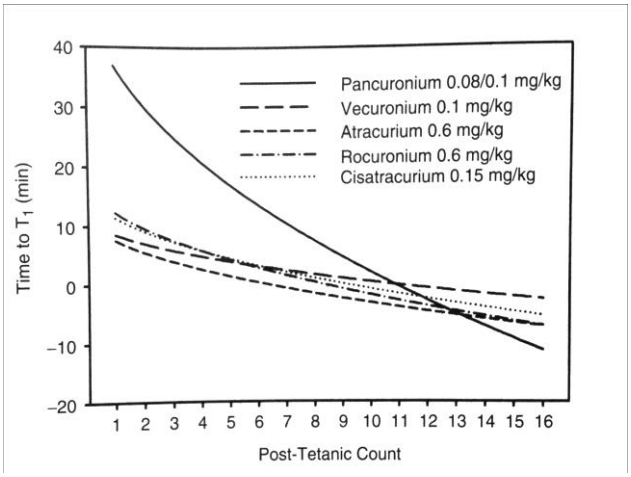
- Patients who have abnormal cholinesterase activity, who have had prolonged exposure to depolarizing agents.
- Also known as mixed or dual type block, they display characteristics of non-depolarizing block i.e. fade in response to TOF or tetanic stimulation and PTP.

Clinical evaluation of responses after administration of Non Depolarising NMBA

Injection of a dose of depolarizing muscle relaxant to allow intubation, results in 3 levels of neuromuscular blockade (NMB), Intense blockade, moderate or surgical blockade and recovery. Different muscle groups require display degrees of relaxation during surgery and recovery e.g. intubation requires intense NMB or abdominal wall relaxation moderate NMB for surgery.



Intense NMB
Time to intubating dose depends on the drug and dose administered.
There will be no repose to single twitch or TOF.
The duration of this period is dependant on patient and drug factors.
The PTC can be used to predict the appearance of the first twitch during this period. Clinically this may be applicable where a large dose of muscle relaxant is given and the case is subsequently cancelled.



Moderate or Surgical Blockade
This begins with the appearance of T1 and ends when T4 appears.

Recovery
The recovery phase starts once T4 appears i.e. less than 75 percent of receptor sites are occupied by NMBA.
When the TOF ratio is greater than 0.4 it becomes very difficult to assess the presence of fade without objective monitoring. The DBS may help in this situation.
Different muscle groups recover at different rates, which may have clinical implications in evaluating residual weakness after administration of NMBA.

What TOF ratio indicates recovery from NMBA?

Previously a TOF ratio of 0.7 was thought to be adequate for sufficient recovery in the postoperative period.
This correlates with clinical signs as below:

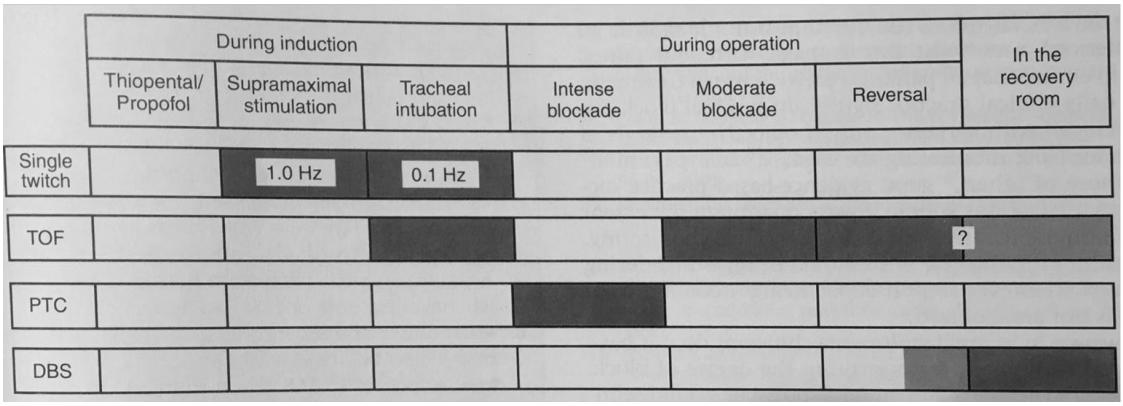
Clinical tests of postoperative muscular recovery

Unreliable
Sustained eye opening
Protrusion of the tongue
Arm lift to opposite shoulder
Normal tidal volume
Normal or near normal vital capacity
Maximum inspiratory pressure <40-50 cm H_2O
Reliable
Sustained head-lift for 5 seconds
Sustained leg lift for 5 seconds
Sustained hand grip for 5 seconds
Sustained "tongue depressor test"
Maximum inspiratory pressure ≥40-50 cm H_2O
(Normal swallowing?)

- TOF ratios below 0.7 are a significant risk factor for postoperative pulmonary complications.
- TOF ratios of between 07-0.9 decrease chemoreceptor sensitivity to hypoxia
- TOF ratios of less than 0.9 are associated with increased risk of regurgitation and aspiration, diplopia, and subjective feelings of weakness.

Residual neuromuscular blockade is inadequate neuromuscular recovery as measured by objective neuromuscular monitoring which may show a TOF ratio of less than *0.9*

What Nerve Stimulator mode to use when ?



? = TOF is less useful in the recovery area unless some form of objective monitoring is used

Applicability to Sugammadex

Sugammadex is a modified gammacyclodextrin that binds aminosteroid NMB (rocuronium>vecuronium>pancuronium). The TOFC may be useful to determine the most appropriate dose to ensure full reversal. At a TOFC count of 2, a dose of 2mg/kg will reliably produce TOF>0.9 in 2 minutes.

Key points:

- Clinical evaluation of recovery of neuromuscular function has limitations.
- Residual neuromuscular weakness has significant morbidity.
- Absence of fade does not exclude significant residual block.
- Objective neuromuscular monitoring is important.
- Clinically significant residual neuromuscular blockade is less likely if the TOF ratio (as measured objectively) is greater than 0.9.

References and further reading:

1. Anaesthesia Monitoring Techniques, Anaesthesia and Analgesia 2002, Hughes, Griffiths

2. Neuromuscular Monitoring. Miller (Ed). Anesthesia 6th ED (pp1551-1569)

3. Physiology of Neuromuscular Transmission and Nerve Stimulators. Wilson , UCT Part 1 ARC , 2006

4. Monitoring Neuromuscular Transmission, Torda, Anaesthesia and Intensive care , Vol 30 , No 2 . 2002

5. Residual Neuromuscular Blockade , Moi, 2013 , WFSA

Understanding Ultrasound

Dr Renier Verbeek

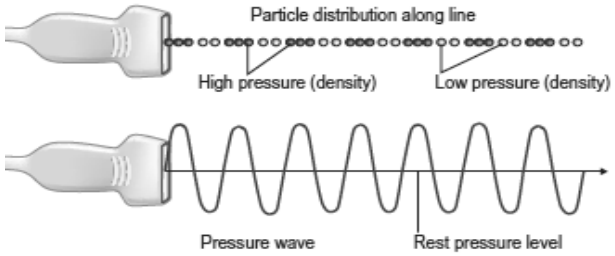
Private Practice & Honorary Lecturer
UCT Department of Anaesthesia & Perioperative Medicine

Introduction

Ultrasound has a continuing expanding role in all spheres of medicine, including anaesthesia: regional anaesthesia, transesophageal and transthoracic echocardiography. It is important to understanding the fundamental physics, in order to appreciate limitations and unwanted artifacts

Sound Waves

Soundwaves consist of longitudinal mechanical waves with regions of high and low pressure, referred to as areas of compression and rarefaction, generated by an oscillating source.
(Rarefaction is the reduction in the density and pressure of a medium)
Sounds waves are transmitted from particle to particle
Particle moves in the same direction as the wave
Separation between high & low pressure areas depends on wavelength (therefore frequency)



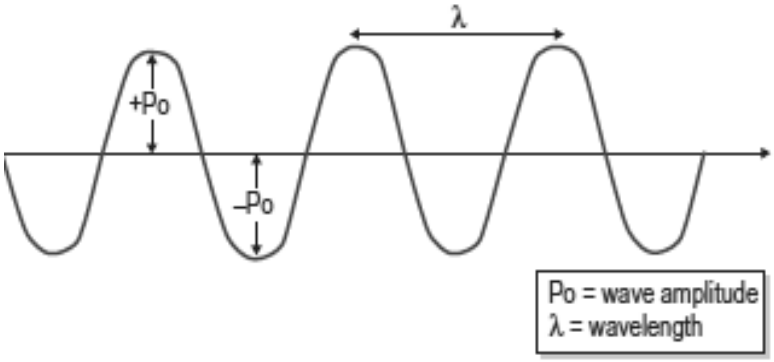
Principles

Wavelength λ
- Distance between crests of a sinusoidal wave
- Distance between points of equal pressure, in phase between 2 consecutive waves

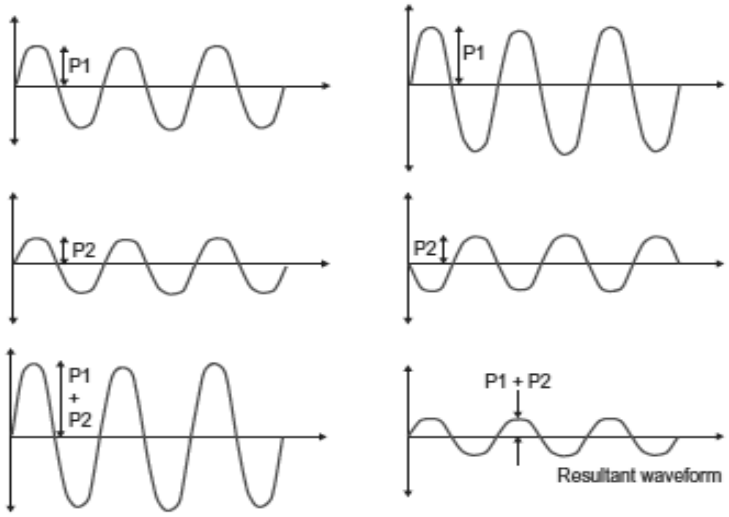
Frequency
- Waves/cycles per second (Hz)

Amplitude
- Intensity of sound
- Measured from baseline to maximum displacement

Pitch
- Determined by frequency

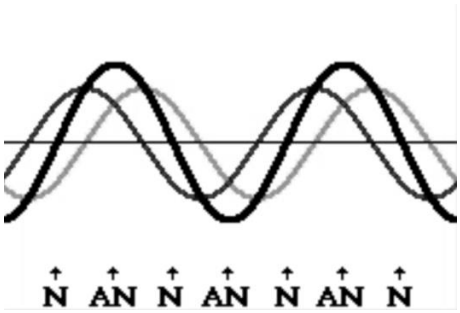


The distance between corresponding points on successive waves defines the period (T).
The reciprocal of the period is the frequency.
Waves with identical frequencies that start simultaneously are in Phase.
Waves that start at different times are Out of Phase, described by the Phase Angle.
When two waves of identical frequency, amplitude and phase combine, a single wave with twice the amplitude results (constructive interference).
If the waves are 180 out of phase they will cancel each other out (destructive interference).



Standing Wave Pattern

Waves of the same frequency travelling in opposite directions interfere
When the distance to the reflected surface is $^1/_2$ **λ** or multiples thereof, specific points appear to be standing still. These points are called '*Nodes*' and '*anti-node*s'. Destructive and Constructive interference occurs in such a way that the points appear to be standing still. A marked increase in the pressure amplitude is observed.



Harmonics

When the separation of the surfaces is one half-wavelength, a strong resonance occurs, known as the fundamental resonance. When separation between 2 surfaces are equal to multiples of the half wavelength, weaker resonances occur known as harmonics. Conversely harmonics can also be seen when the distance is constant and the wavelength is varied. These harmonics occur when the wavelength equals some multiple of the separation.
These frequencies are known as harmonic frequencies. They are only created within a medium at specific frequencies of vibration. At any frequency other than a harmonic frequency, the interference of reflected and incident waves leads to a resulting disturbance of the medium that is irregular and non-repeating.

Ultrasound

Infrasound 0-20 Hz
Audible (acoustic) 20-20000 Hz
Ultrasound > 20000 Hz
Medical use 20 000 – 50 000 000 Hz (20 kHz – 50 MHz)

Propagation Velocity

C (Propagation Velocity) = λ (wavelength) x f (frequency)

Velocity depends on the density $\boldsymbol{\rho}$ and stiffness $\boldsymbol{\beta}$ of the medium

$$c = \sqrt{\frac{\beta}{\rho}}$$

Velocity is inversely related to the density and directly related to stiffness. Density of the medium is related to its weight and the stiffness of the medium is related to its "squishability". As the medium becomes more dense, the slower is speed of ultrasound in that medium. The stiffer the tissue, the faster the propagation velocity (direct relationship). The more rigid the material, the higher the speed.
Most machines use an average speed of 1540m/s.
This is independent from frequency over a wide range 1 – 50Hz.

Safety

TI refers to the thermal index of the transducer. The potential to cause tissue heating
MI refers to the mechanical index of the transducer. The potential to cause mechanical effects, like cavitation, causing microbubbles.
Streaming, which is fluid movement at a microscopic level.

Velocity and Temperature

In an ideal gas, the speed of sound is proportional to the square root of the temperature
Ultrasound has a fixed velocity in a specific medium
1540 m/s in tissue at $37^{o}C$
344 m/s in air at room temp.
1450m/s in water @ $15^{o}C$

Acoustic Impedance

Density = mass/volume

ρ = m/V (g/cm^3)

Acoustic impedance = ρ ($g/cm^{3)}$ x C (cm/s)

Acoustic impedance measured in Rayls. $g.cm^{-2}.s$

Amplitude

Difference between the peak value and the average value of the waveform
Expressed in decibels or dB

Decibels dB

"Effective sound pressure relative to the threshold of hearing at 1 kHz on a logarithmic scale"

Amplitude determines the intensity
Amplitude decreases usually by 1 dB per 1 MHz per 1 centimeter traveled

Intensity

"Rate of energy flow per unit area" or the Power per unit area

$$\frac{W}{m^2}$$

Intensity is expressed as effective sound pressure relative to the threshold of hearing at 1 kHz. In dB. Logarithmic scale. intensity is normally measured with a hydrophone, which takes the form of a small probe with a piezoelectric element at its tip. Usually measured at the focus of the field or 1-2cm from transducer I_{spta} refers to Intensity at spatial peak averaged over time, at the focus point.

I_o = maximum intensity of the machine
The decibel notation relates the current Intensity I to a reference value I_o
Output Intensity (dB) = 10 Log (I/I_o)

Power

Power is the rate of energy flow through the cross-sectional area of the ultrasound beam
The ultrasound field or beam is the region in front of the transducer that is affected by the transmitted vibration. The Power Control button was historically described i.t.o. decibels in older ultrasound machines. The Power control selection reflects the ratio of the Reference and Current power.

Output power in decibels (dB) = 10 log (P/P_o)

Power is the rate of energy flow (W)

Power can be measured with a radiation power balance. The force exerted on an object placed in water, in an ultrasound field is directly related to the power of the ultrasound field

Radiation force = $2P\ c$

The force experienced is 0.135 mg per milliwatt. Power levels down to 0.1 mW can be measured

Tissue Interactions

- Reflection
  - Specular
  - Scattering
- Diffraction
- Refraction
- Absorption
- Echo (Reflection)

} Attenuation

Waves interact within tissue and between tissues with a different acoustic impedance. They can bounce (reflection), bend (refraction), scatter, diverge(diffraction), and convert their kinetic energy into heat energy (absorption)

Soundwaves are partly transmitted and partly reflected at the interface between different tissues
A sudden change in the density of the substance results in reflection
The larger the change in acoustic impedance the greater the reflection and will be highest at gas/tissue interfaces (Importance of gel)
The angle of the beam to the structure influences reflection/refraction

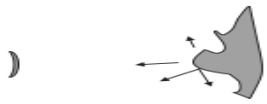
Reflection

Specular reflection occurs when the wave is reflected off a surface, where the length of the structure is significantly longer than the wavelength of the ultrasound. If the reflector is smooth and the ultrasound beam is perpendicular to the surface, then the reflection is strong and called specular. If the incidence angle is not perpendicular, then specular reflectors are not well seen
Eg. Blood vessel, tendons, tissue boundaries

Scattering

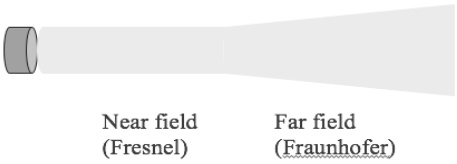
If the reflector is much *smaller than the wavelength* of the ultrasound, the ultrasound is uniformly scattered in all directions and this is called **Rayleigh scattering**, where the reflection is equal in all directions.

When the ultrasound wavelength is almost equal to tissue interfaces, we get **DIFFUSE REFLECTION**
At the irregularities of the boundary, the ultrasound is redirected in all directions or scatters. The echo partially returns to the transducer and is used to produce the image. These small reflectors give the texture of the organ being insonated



Diffraction

Spreading out of a wave as it passes from its source through a medium
The pattern is dependent on the shape and size of the source relative to the **λ**
To create a parallel beam, the diameter of the crystal face should be > 10 times that of the ultrasound wavelength
A Disc-shaped crystal is often used in medical U/S. The ultrasound beam will be cylindrical for a short distance, after which it diverges at a small angle



The diffraction pattern from a small transducer is divergent at a larger angle from close to the transducer
Diffraction also occurs beyond an obstacle such as a slit aperture (Between ribs)
The sharpness of a beam is ultimately determined by diffraction.
The higher the frequency, the narrower the beam and the finer the image detail.

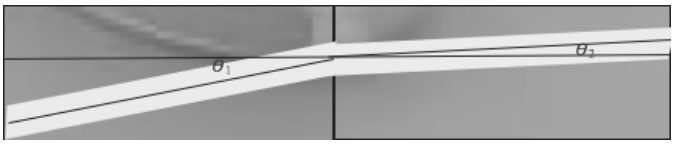
Refraction

"Change in direction of a transmitted wave between two media with differences in acoustic impedance at an incidence angle other than 90^{o}"

Snell's Law

$$\frac{\sin \theta_1}{V_1} = \frac{\sin \theta_2}{V_2}$$

Snell's Law describes the relationship between the angles and the velocities of the waves. Snell's law equates the ratio of material velocities V1 and V2 to the ratio of the sine's of incident (θ_1) and refracted (θ_2) angles



Angle of refraction usually insignificant, but degrades the image and can result in an error in location of image

Angle of refraction in specific tissues

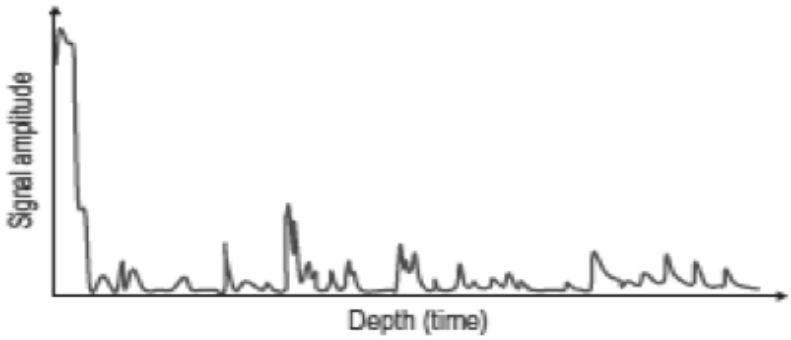
Bone/Soft tissue	19°
Muscle/Fat	2.5°
Muscle/Blood	0.5°
Muscle/Water	1°
Lens/Aqueous humor	2°

Absorption

As an ultrasound wave passes through tissue, its energy is converted into heat energy, the pressure amplitude reduces with distance travelled. The higher the frequency, the more rapid the amplitude reduction. The type of tissue also influences the absorption rate

Attenuation

Image information is provided by the energy of the reflected waves. The depth of each reflected wave is determined by the time taken for the wave to travel from the transducer to the tissue interface and back. The Ultrasound waves undergo ***attenuation***. The Intensity decreases exponentially with depth
Tissue penetrance is ***inversely related to probe frequency***



.

The Doppler Effect

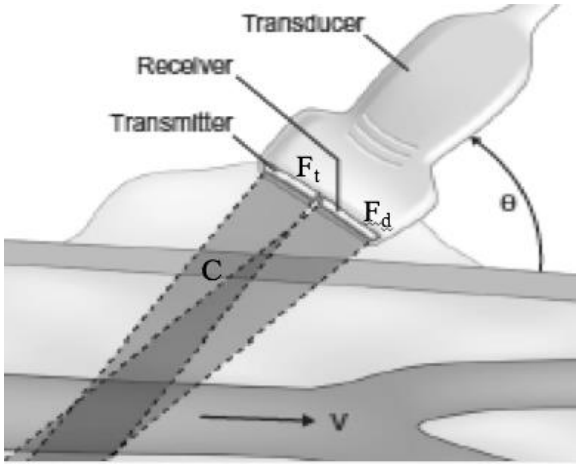
In 1842, Johann Christian Doppler (Austrian Physicist), published his most famous paper “Über das farbige Licht der Doppelsterne” (“Concerning the Colored Light of Double Stars”), which contained his first statement of the Doppler effect:

‘There is an apparent change in the frequency or wavelength of a wave when there is relative motion between the source of the waves and an observer’

Doppler Equation

$$F_d = \frac{2 F_t V \cos\theta}{C}$$

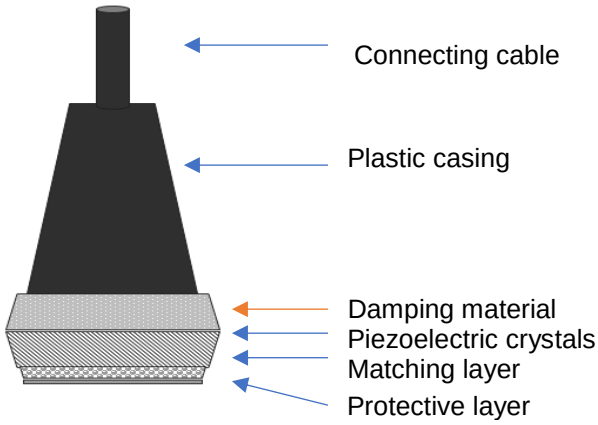
F_d = Doppler shift
F_t = Transmitted doppler frequency
$\cos\theta$ = Cosine of bloodflow to beam angle
C = Velocity of sound in tissue
V = Velocity of blood flow



Producing Ultrasound

Transmitting transducer changes electrical energy into mechanical energy (ultrasound wave)
The wave is transmitted through a medium and reflected from tissue interfaces. A receiving transducer changes the ultrasound wave into an electrical signal, squeezing and stretching of material generates a potential difference. An image is formed on the monitor.

Transducer



The Piezoelectric Effect

The ***direct piezoelectric effect*** refers to a change in electric polarization that is produced in certain materials when they are subjected to mechanical stresses.
The ***inverse piezoelectric effect*** refers to a deformation of these materials that results from the application of an electric field

Piezoelectric crystals

Artificial polycrystalline ferroelectric materials like lead zirconate titanate (PZT), synthesized. Engineered materials include lithium niobate and lead zirconate titanate (PZT). Heating the PZT past its Curie temperature 3280 – 3650 ^{o}C and applying an external voltage, aligns the dipoles in the ceramic, which is retained when it has cooled. The crystalline material is then cut into shape
Naturally available crystalline materials, include quartz, potassium sodium tartate (Rochelle salt)
Thickness of the crystal determines resonant frequency

Producing ultrasound: A voltage is applied across piezoelectric material which causes a 3D change, resulting in expansion & contraction. If the rate of compression and rarefaction exceeds 20000 times/sec, we have ultrasound.
The piezoelectric crystal will resonate at its own frequency until the voltage is discontinued, after which the resonance will decay.
A Pulsed Wave is produced when a voltage is supplied for an extremely short period; Continuous Wave is produced when a voltage is maintained.

Pulsed Ultrasound

Pulse Duration = the time that the pulse is on.

It is determined by the number of cycles and the period of each cycle. In clinical imaging, a pulse is comprised of 2-4 cycles and the pulse duration is usually between 0.5 to 3 microseconds

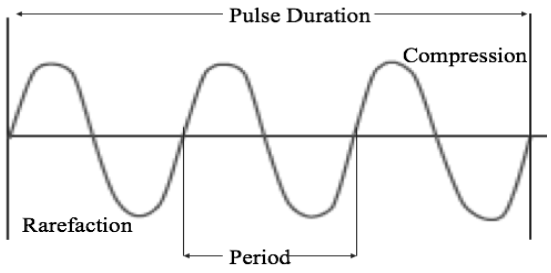
Pulse Duration (msec) = #cycles x Period

$$\lambda \text{ (mm)} = \frac{C \text{ (mm/msec)}}{F_t \text{(Hz)}}$$

$$\frac{1}{F_t} = \frac{\lambda}{C} \quad \text{and Period} = \frac{1}{F_t}$$

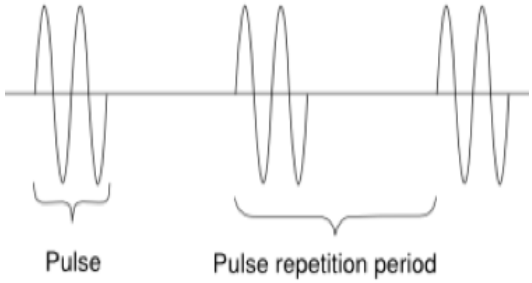
Pulse Duration = #cycles x λ

C



Pulse Repetition Period

Time between the onset of one pulse till the onset of the next pulse
Changing the depth, changes the PRP



Pulse Repetition Frequency is the number of pulses that occur in 1 second

Duty Factor (DF)

DF is the percent of time that the ultrasound system is on while transmitting a pulse
DF in clinical imaging is usually between 0.1% to 1%, so the machine is mostly listening!
An image cannot be created with a continuous wave

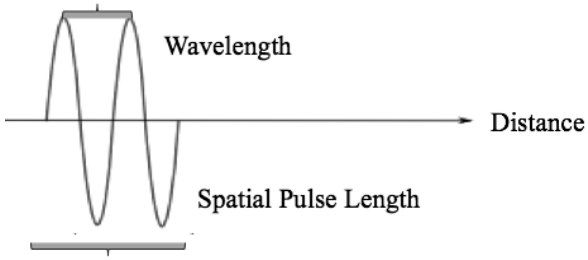
$$\frac{\text{pulse duration (sec)}}{\text{pulse repetition period (sec)}} \times 100$$

Spatial Pulse Length

Spatial Pulse Length is the distance from the beginning of one pulse till the end of that same pulse.
Typical values range from 0.1 to 1 mm

SPL (mm) = # cycles x wavelength (mm)

Axial resolution = SPL/2



Focus

A beam has two distinct regions. In the near field (Fresnel), the beam has a constant diameter, determined by the diameter of the transducer. The length of the near field is related to the diameter, **D**, of the transducer and the wavelength, **λ**, of the ultrasound by

$$\textbf{Near field length} = D^2/4\lambda.$$

The wavelength is inversely related to frequency. Therefore, for a given transducer size, the length of the near field is proportional to frequency.

In the far field (Fraunhofer), the beam diverges. This causes the ultrasound pulses to be larger in diameter, with less intensity along the central axis. The approximate angle of divergence is related to the diameter of the transducer, **D**, and the wavelength, **λ,** by

Divergence angle (degrees) = 70λ /D.

Divergence is decreased by increasing frequency. The major advantage of using the higher ultrasound frequencies (shorter wavelengths) is that the beams are less divergent and produce better detail, but they are depth limited.

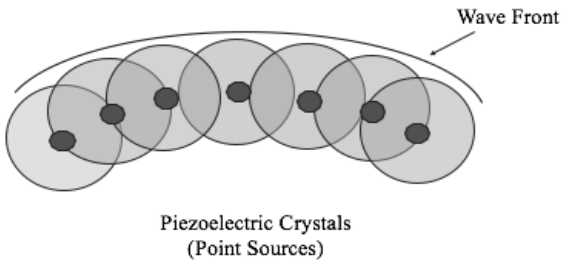
Focusing

A transducer can be designed to produce a focused ultrasound beam by using a concaved piezoelectric element or an acoustic lens in front of the element.
Weak focusing produces a long focal zone and greater depth. A strongly focused transducer will have a short focal zone and a shorter focal depth. The focus of these transducers are usually fixed and are unable to accommodate different depths

Adjustable transducers have multiple crystal elements of the same frequency arranged in Linear, Curved or Annular Array

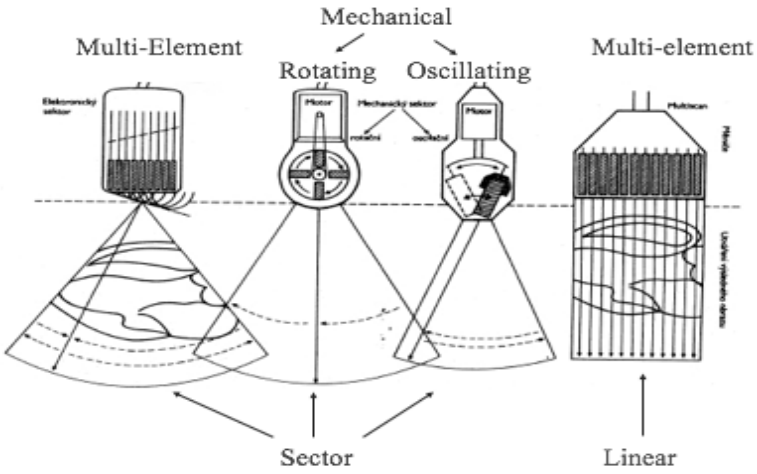
Ultrasound Beam

Each probe has multiple crystals that are electronically and acoustically isolated. A single wave front is formed (Huygen's Principle). In 1678, the Dutch physicist Christian Huygens (1629- 1695) stated that the wave front of a propagating wave of light at any instant conforms to the envelope of spherical wavelets emanating from every point on the wave front at the prior instant".
Applying this to ultrasound waves, "The wave front is the sum (or envelope) of the individual wavefronts from each piezoelectric crystal".



Transducers

There are 2 types according to the image created.
A sector image can be produced by a flat or curved probe. This provides information from a large area.
Linear images are usually obtained from a flat or linear probe. A small area is scanned, with good image resolution



Mechanical Transducers

Mechanical transducers contain 1-5 crystals that are mechanically moved either in a rotational or oscillating way. Focusing of the ultrasound beam is accomplished with a Lens or a Concave crystal, this results in very limited focal control. The focal zone has an increased energy concentration, but unfortunately there is a limited frame rate and they are unreliable.

Multi-Element Transducers

An Array is a group of piezoelectric crystals within the transducer housing

Linear Sequential Array
Curved Sequential Array
Linear Phased Array
Curved phased Array
Annular Array
Multi D Array

Linear Sequential Array (rectangular image)

120-250 Crystals (Each crystal has 1xλ width)
Arranged in a flat line
Activated in groups (eg.1-5, 6-10…),
Beams are parallel
Focused originally with lenses, now electronic
Image corresponds to the width of the actual transducer
Linear array produces a rectangular image, ideal for vessels, vascular access, needle guidance.
It has a wide near field, but limited depth

Curved Sequential Array (Sector image)

120-250 crystals (1xλ width)
crystals are arranged along a curved surface
Activated in groups
Fan shaped image
Electronically focused
Curved array produces a wide near and far field. Used in abdominal and obstetric imaging.
Good depth penetration, but there is a loss of lateral resolution with increasing depth

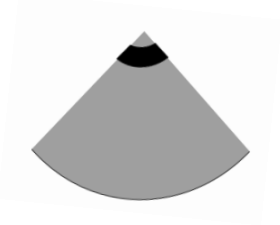
Linear Phased Array

100-300 crystals
The crystal have a width of ½ - $^{1}/_{4}$ λ
Crystals fired at different times with an extremely small delay
Fan shaped image
Beam can be steered and focused
Small 'footprint' allows for imaging between ribs and in small places

Annular Array

Single crystal cut into rings
Mechanical beam steering
Outer diameter crystal has a deeper focus point
Smaller diameter rings have a shallow focus point
Sector or fan shaped image
A 6-ring element transmits 6 pulses, one for each focal depth
If 1 ring is damaged, a horizontal image loss at that focal depth

Annular ARRAY



SECTOR IMAGE
WITH DAMAGED
CRYSTAL

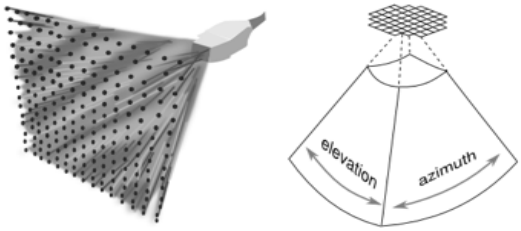
Dynamic Receive Focus

Activation of crystals can also be controlled in receiving mode. By applying a time delay before activation, the transducer can have a higher sensitivity for echoes coming from a specific depth. The receiving focal depth can be changed rapidly and the focus can be swept through a range of depths to pick up the multiple echoes produced by one transmitted pulse.

Multi D Array

In a multi D Array, the piezoelectric elements are arranged in a 2D matrix. Each PZ element represents a scan line, by combining all the data, a 3D set is reconstructed. With careful timing of each crystal, a pyramidal volumetric data set is created. When imaged several times per minute (>20), a real-time image is achieved.

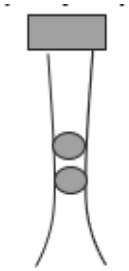
Multi D Array



Resolution

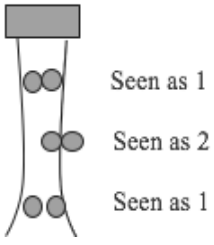
Axial

Along axis of beam
Discriminating between 2 separate objects parallel to the beam
Affected by frequency and pulse length. The higher the frequency and the shorter the pulse length, the better the axial resolution



Lateral

2-point discrimination perpendicular to the beam
Best at the focal zone. Pulse width usually larger than pulse length, therefore the lateral resolution is usually worse than axial resolution



Spatial

Combination of above

Contrast

Discriminating between 2 objects/dots of similar intensity
differentiation between tissues having different characteristics e.g. liver/spleen.

Temporal

Temporal resolution is the time from the beginning of one displayed frame to the next
It represents the ability of the ultrasound system to distinguish between instantaneous events of rapidly moving structures, for example, during the cardiac cycle. Anatomical structures are displayed as sequential frames over time.
Frame rate is limited by the frequency and depth
The larger the depth, the slower the frame rate
The higher the frequency the higher the frame rate
M-Mode has the highest temporal resolution, because it has only 1 scan line

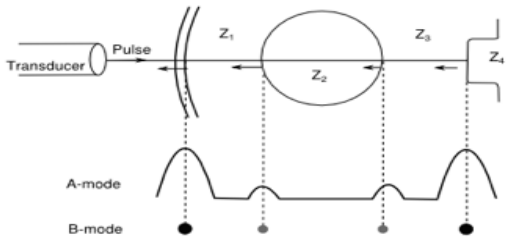
<u>Imaging Modes</u>

A-Mode

Amplitudes of the returned echoes are displayed
Spikes are produced when a single beam passes through different surfaces/objects
The distance between spikes is the speed of u/sound divided by half the time
Historical

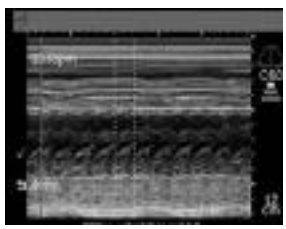
B-Mode

Two dimensional
Brightness corresponds to amplitude
White (bright) dots: bone, gallstones
Grey dots: solid organs and thick fluid
No reflection: fluid/blood/urine



M-Mode

Single beam. M-Mode displays the depth of reflected echoes in a single scan line on the y axis in relation to time on the x axis, with pixel brightness corresponding to the amplitude
High sampling frequency (1000/s). Movement of a structure and rate of motion eg. Valve, ventricular wall



2D

Multiple crystals – linear, curved, phased-array or a Moving crystal
Sequential B-mode pulses
2D B-mode most common
    Assess motion and anatomy
    Rapid series of 2D image

The reflected wave is represented by a dot. The position represents depth
Depth is proportional to the time delay between the reflected and transmitted wave

$$\textit{Distance} = \frac{\textit{Velocity}}{\textit{time}}$$

The brightness represents the amplitude of the dot
Amplitude determines the intensity (loudness)
(Intensity = Watts per square meter)
Amplitude is related to the energy of the reflected wave
Combination of dots form image (pixels)
Resolution relates to number of dots formed
The higher the frequency the better the resolution
Lower frequencies penetrate deeper (poorer image quality)

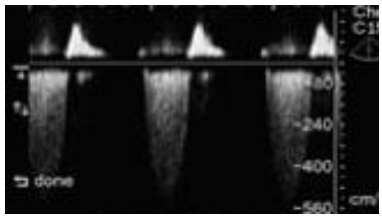
Doppler

Pulsed Wave

The same crystal sends and receives
Short pulses of sound
Depth precision

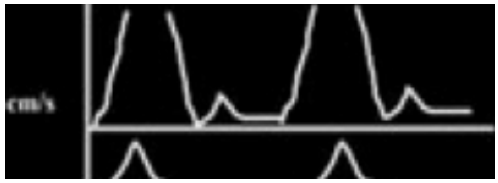
Continuous Wave

Separate sending and receiving crystal
No depth precision. It displays all velocities in the scan line
Used when interrogating high velocity jets



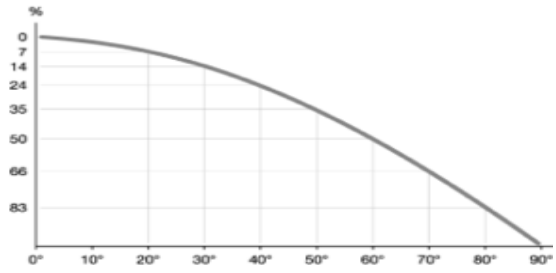
Aliasing

Aliasing occur when the sample rate is not adequate for high frequencies eg. High velocity flow
Peaks are cut off and displayed below the baseline



Spectral Waveform

A spectral waveform gives an audible signal; the density of the line is related to the quality of flow. The y axis depicts the flow velocity in cm/s. The x axis is time. The direction of flow is towards the transducer above the baseline and away from the transducer below the baseline
Errors will occur when the Doppler angle exceeds 20 degrees



Percentage underestimation on y axis as a function of the angle of insonation on the x axis

Colour Doppler

Pulsed Wave is used. The colour is superimposed on a 2D image
Red/blue provides info regarding direction of flow. The top bar is flow towards and bottom is flow away
Prone to aliasing with high flows (high frequencies). A useful mnemonic is BART (Blue Away and Red Towards)

CPA

Colour Power Angiography
Amplitude of blood cell motion
Superimposed on a 2D picture
Single continuous colour
No direction
Less angle dependent
Ideal in low flow states

Duplex
Combines Doppler and image

Optimizing Image

Gain

Gain controls the brightness of the image. There is depth dependent attenuation of signal
An incorrect gain setting can influence the quality of the image.
Insufficient
- Misses low reflecting structures

Excessive
- False echoes/oversaturation
- Obscures important diagnostic characteristics

Near, far (time gain) and overall gain settings
Gain is expressed in terms of Volts or Decibels
Output Volts/Input Volts
Gain (dB) = 20 log (V_{out} /V_{in})

Depth

A tradeoff between visualizing detail vs. visualizing relevant structures

A lower frame rate is required for deeper structures (reflected waves take longer to reach probe)
Too shallow results in the inability to image deeper structures
The target structure must be centered in the middle towards to bottom 2/3 of the image

Focus

Focal zone
Depth and size can be adjusted

Correct transducer

Frame averaging

Eliminates noise

Edge enhancement

Increases visual perception using specific software

Gel

Required to eliminate surface air (large difference in density will result in excessive reflection)
Should cover scanhead completely

Transducer Orientation

Transducer notch should point to right side in transverse views
Head in longitudinal view
Coloured orientation mark on screen

Artifacts

Some artifacts are ***technique dependent*** and some are ***sonographic*** due to the properties of ultrasound waves and some are a combination thereof.

Technique

Excessive brightness or noise caused by excessive gain setting
Low signal due to inappropriate probe selection or low gain setting
Anisotropy
Image dependent on the angle of the beam. 90^{o} good image, $< 60^{o}$ no image

Sonographic artifacts

Acoustic Shadowing
Dark area behind a strongly reflective tissue (eg. rib)

Acoustic Enhancement
Sound travelling through fluid filled structures without attenuation (post-cystic enhancement).
Decrease far gain, reduce overall gain

Reverberation
Interface between 2 markedly different acoustic structures. Multiple lines seen on image. Ultrasound wave reflects back and forth. The ultrasound transducer interprets the sound waves returning from the reverberation as deeper structures since it took longer for the wave to return to the transducer.
Change transducer angle, Apply gel, Rotate patient

Mirror Image artifact
The primary beam reflects from a highly reflective surface (e.g. diaphragm) but instead of directly being received by the transducer, it encounters another structure (e.g. a nodular lesion) in its path and is reflected back to the highly reflective surface (e.g. diaphragm). It then again reflects back towards the transducer.
The ultrasound machine makes a false assumption that the returning echo has been reflected once and hence the delayed echoes are judged as if being returned from a deeper structure, thus giving a mirror artifact on the other side of the reflective surface.

Comet-tail artefact is a grey-scale ultrasound finding seen when small highly reflective objects are interrogated and is a basically a form of reverberation artifact.

Beam width artifact

Side lobe is reflected ultrasound from a strong reflector that is outside of the central beam, and where the echoes are displayed as if they originated from within the central beam. Secondary beams occur because the crystals also expand and contract radially. These radial beams are called ***side lobe*** beams. Side lobe beams are low intensity beams that surround the central beam.
Side lobe artifacts are echogenic, linear or curvilinear artifacts. Strong reflectors include bowel gas adjacent to the gallbladder or urinary bladder.

Grating lobe

Velocity Artifact
Inaccurate depth. Displacement of an echo deeper than its actual position in an imaged structure. The artifact is due to the image processor's assumption that the velocity of the ultrasound beam within the imaged anatomy is uniform

Refraction effects. The transmitted pulse strikes an interface at a non-perpendicular angle. The difference in propagation speeds between the two tissues can cause refraction to occur. Should the refracted incident sound wave strike a reflector and cause an echo to return to the transducer, this may be displayed at an incorrect location as the transducer assumes all echoes have traveled along a direct path.

Advances in Imaging

Fundamental frequency.
Natural frequency at which a substance will resonate/produce sound

Harmonic
Multiples of the fundamental frequencies

Band width
is the range of frequencies within a pulse and is the difference between the highest and the lowest frequencies

As ultrasound travels through tissue, it has a non-linear behavior. In waves with large pressure amplitude, the speed of sound is higher in regions where the pressure amplitude is positive than it is in the regions where the pressure is negative. The speed is different in the regions experiencing positive half-cycles than in negative half-cycles because the density of the medium changes with pressure. The effect is such that as the waveform passes through the tissue, the positive half-cycles catch up on the negative half-cycles resulting in distortion
Some of its energy is converted to frequency that is doubled (or second harmonic) from the initial frequency that is used (or fundamental frequency). The deeper the target the more ***second harmonic*** frequency is returned. As the ultrasound beam travels through tissue, new frequencies appear that can be interrogated. *Second harmonic data gets less distortion*; thus, it produces better picture and the second harmonic is strongest in the *center of the beam*, thus it has less side lobe artifacts
Deeper tissue resonates at twice the fundamental frequency (^{2nd} harmonic).
Extended resolution harmonics enables a reduction in image artifacts

Sonographic appearance of Structures

Arteries	Anechoic, pulsatile
Veins	Anechoic, compressible
Tendons	Long: tubular, loosely packed, blurred lines Transv: Circular, pale halo
Nerves	Long: Tubular, bright surface, bright inner lines Transv:Circular, bright surface, hypoechoic black dots, bright oulines
Pleura	Hyperechoic lines
Lung	Multiple grey echoes
Bone	Hyperechoic

Bibliography

1. Bosenberg, A. Principles of ultrasound for the Anaesthesiologist
2. Hatfield, A. et al. Ultrasound for central venous access. CEACCP Vol. 5. No. 6.2005
3. Carty, S. et al. Ultrasound guided regional anaesthesia. CEACCP Vol. 7. No. 1. 2007
4. AnaesthesiaUK
5. www.nesra.co.uk
6. Sprawls, P. Ultrasound Production and Interaction
7. Basic Principles of Ultrasound. Dcconferences.com.au
8. McDicken, WN et ak. Basic Physics of Medical Ultrasound

Gas and Vapour Analysis in Anaesthesia

Dean Nolte

Private Practice & Honorary Lecturer
UCT Dept of Anaesthesia & Perioperative Medicine

Gas analysis in anaesthesia practice serves several important purposes/ functions. These include safety (determining the anaesthetic agent being used, percentage of oxygen being administered, as well as the analysis of any harmful gases – H_2 and CH_4), assessment of the physiological impact of the anaesthetic agents (cardiovascular and respiratory), and is useful in laboratory research. As a broader (societal measure) these techniques of gas analysis are employed to measure the presence of atmospheric pollution. This is in- and out-of-hospital pollution.

The use of oxygen and carbon dioxide analysers are recommended as a standard of care (AAGBI Recommendations, 2007[1]), and vapour analysis is deemed the basic standard of care when a volatile anaesthetic technique is utilized.

Many techniques employed have become obsolete or are not clinically-practical. These can be subdivided into chemical and physical means of determining gas concentrations. The most useful (in an anaesthetic context) would be **cheap**, easy-to-use (i.e. not cumbersome), have a **rapid** response time, provide **accurate** data from a **continuous** sample line.

Chemical	**Physical**	
	Non-Specific	**Specific**
Colorimetric ∗ Only CO_2 ∗ Use when capnography unavailable	Density	***Electro-magnetic*** ∗ *Infrared absorption* ∗ Ultraviolet absorption ∗ *Raman scattering* (electro-magnetic interaction) ∗ Electro-magnetic emission ∗ Radioactive emission
Haldane apparatus (antiquated)	Viscosity	***Atomic/ Nuclear properties*** ∗ *Mass spectrometer*
	Thermal conductivity	***Magnetic susceptibility*** ∗ Paramagnetic (O_2)
	Sound Velocity	Solute-solvent interaction
	Refractive index ∗ Due to interference e.g. Riken gas indicator	Chemico-physical partitioning ∗ *Piezo-electric crystals* ∗ Gas chromatography
	Solubility e.g. Dräger Narkotest Halothane Analyser	***Electro-chemical interaction (O_2/ CO_2)***

1) Electro-Magnetic (Specific Physical Measurement)

1) Infrared Absorption

This is the most common method of measurement utilized to measure gaseous CO2 in the clinical environment. It is also the most commonly-used method of analysis when a side-stream sampling device is employed.

The principle of measurement derives from the fact that molecules with dissimilar atoms (non-uniform) will absorb infrared radiation at specific frequencies (based on interatomic bond properties). This is converted into energy, and causes molecular vibration. The absorption by the molecule is at different wavelengths. These are specific, and allow identification of the molecule under investigation. The Beer-Lambert Law dictates this investigation.

Beer-Lambert Law

Definition: logarithmic dependence between the transmission of light through a substance, and the concentration of that substance in the medium.

Infrared radiation is focused through a chopper wheel with narrow-band fibres. This allows the selection of the specific infrared wavelengths.

This is compared to initially pulsing the radiation applied which will amplify the signal (as it is transmitted through a microphone).

The sample requires a reference/ baseline sample which allows processing and detecting of the infrared radiation.

The overall signal requires amplification.

What: polyatomic, asymmetric molecules.

- CO_2
- Volatile anaesthesia
- N_2O

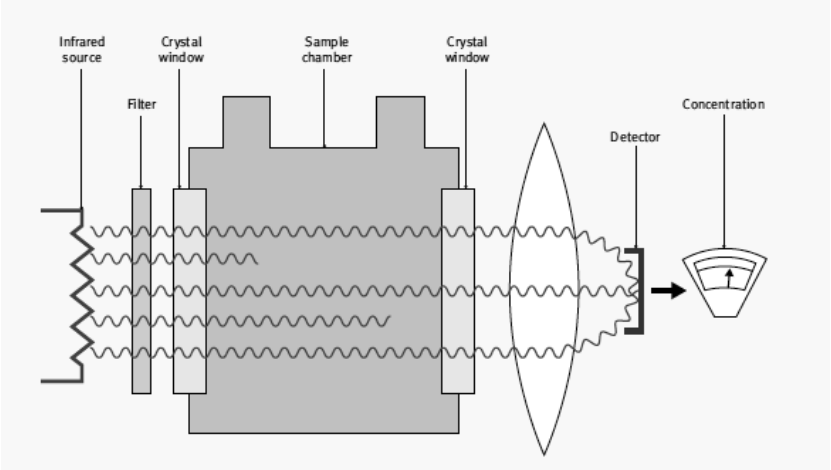
NOT N_2 or O_2

Interference: Alcohol

Water vapour

Carbon monoxide

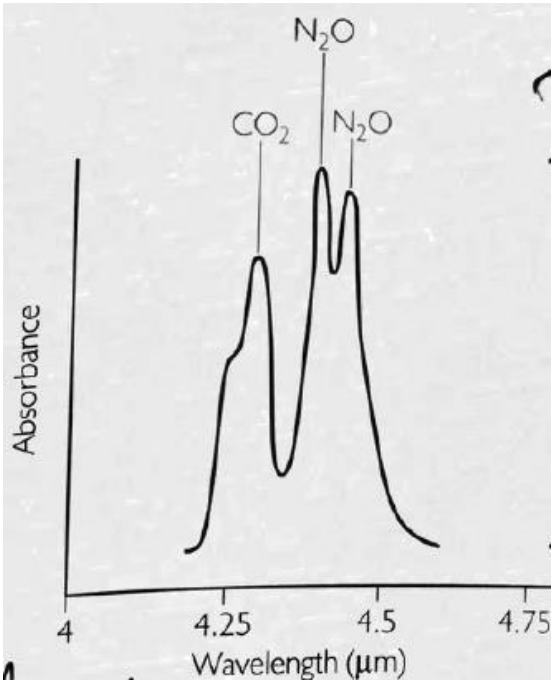
Infrared absorption will aid with analysis of wavelengths of between 3 – 12 μm. absorption is directly proportional to the number of molecules struck by the beam. The increase in the absorption of infrared radiation by bonds of the molecule results in decreased radiation monitored by the detector.



Goswani A, Cooper R. Measurement of gases (O_2, CO_2, N_2, N_2O and volatile agents).[2] Adapted from: Sykes MK, Vickers MD, Hull CJ. Principles of measurement and monitoring in anaesthesia and intensive care. 3rd edn. Oxford: Blackwell, 1991. Figure 17.15, p. 236.

CO_2 and N_2O cause "*collision broadening*". They broaden each other's wavelengths, as well as the peaks of the volatile agents. This energy transfer is compensated for by the electric module.

Infrared	**Side-stream**	**Main-stream**
Setup	Diverting Delayed response (away from patient) ∗ = 2.5 sec/ 3 m of tubing ∗ sample at 150 – 200 ml/min Water vapour can enter the chamber	Non-diverting Shine infrared light through plastic housing, impacting onto a photo-detector. NO gas is sampled from the circuit. BULKY
Volume	150 – 200 ml ∗ Therefore, a problem with very low V_T or low fresh gas flow.	Minimal BUT increase by up to 20 ml of V_D



Davis P, Kenny G. Absorption of infrared radiation by carbon dioxide and nitrous oxide depends on the wavelength of the radiation. In: Basic Physics and Measurement in Anaesthesia. 2005; Fig. 19.4, p.215.

The infrared radiation detector is a lead selenide photocell. There is a pressure change effected which is detected on a diaphragm, and amplified by a microphone (photoacoustic amplifier). This is converted to an electronic signal and displayed on the monitor.

Problems highlighted:

- Collision broadening
- H_2O molecules/ vapour trapped
- Absorption wavelength overlap
- Frequent calibration required against known sample (this does tend to happen at inopportune times)

Infrared absorption does have major benefits and are commonly used mainly due to accuracy of sample measurement, low cost, rapid analysis, and small size (not bulky).

2) Raman Spectroscopy

This is a form of molecular spectroscopy. Electromagnetic radiation is applied from a high-energy monochromatic light source.

Laser light source	**Wavelength emitted**
Argon Laser	488 nm
Neon-Helium Laser	633 nm

Which gases?

Any di-/ polyatomic molecules (this includes those NOT asymmetrical).

- O_2
- N_2
- CO_2
- N_2O
- Volatiles

Mechanism

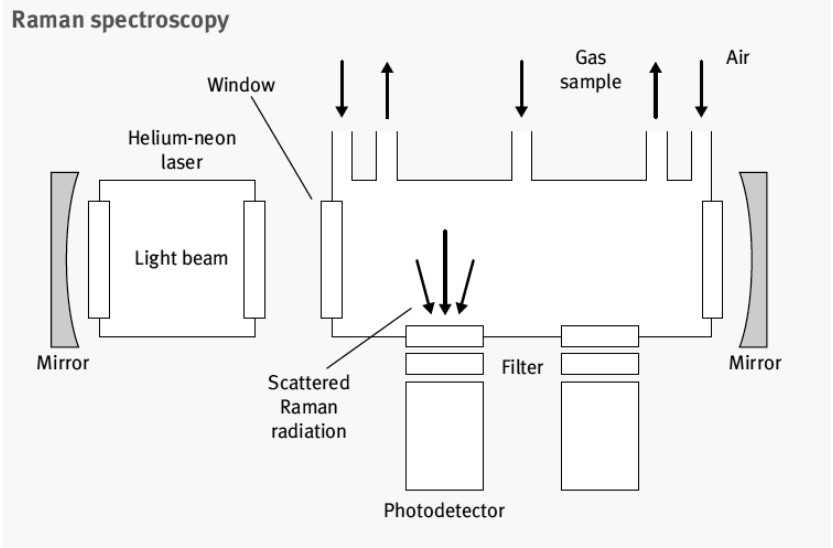
Light energy applied to a medium will create scattering of the light beam. The proportion of this **scattering** can be determined.

	Rayleigh Scattering	**Raman Scattering**
Definition	Most (>99%) of monochromatic light impacting a polyatomic gas molecule will be absorbed and re-emitted in the same direction as the wavelength of the incident beam.	$1/10^6$ of molecules will lose energy due to their interaction with the applied light beam. This is due to absorption and scattering of the molecules by the light beam. Thus, the molecules are excited to higher vibrational levels, and the light will travel at different wavelengths. (It is important to note that this energy transfer is specific to the type of molecule under investigation: the vibrational energy between two molecules in a bond is molecule-/ bond-specific.)
Scattering	Elastic scattering	Non-elastic scattering
Examples	Blue sky, sunsets	Anaesthetic gas analysis, Hard-disk drives, geology

The ***Raman scatterin***g is of low intensity. The applied light source is intense (from a laser), attempting to achieve as much radiation/ Raman energy scattering as possible. The proportion of scattering is analysed perpendicular to the incident light beam: low-intensity Raman light would not be identified amidst the powerful laser beam.

Eight detectors (with filters) are found opposite to the gas sample. These measure a specific shifted frequency (i.e. ↓ f, ↑ λ) for each component gas. There is a -C-H- bond detector which can determine the possible presence of a volatile agent, and then there is a volatile-specific detector **which** will detect which volatile is present.

The concentration of the photons emitted is directly related to the concentration of the vapour. The changes of incident radiation will be transformed into an electrical signal (for measurement and processing).



Advantages	**Disadvantages**
Multiple agents analysed, all with simultaneous input	Laser technology, and laser safety
The sample gas is unaltered, and returned to the breathing circuit	Very large electricity demand (power the laser beam)
	Large, cumbersome size
	Laser CAN degrade volatiles (but air is blown past cell windows)

Davis P, Kenny G. Basic physics and measurement in anaesthesia. 5th edition. Butterworth: Heinemann, 2005. Figure 20.8, p.225.

2) Atomic/ Nuclear Properties

Mass Spectrometry

This method determines the rate of ionization of particles, and then the subsequent mass:charge (ratio). The magnetic field will thus exert a turning force on an electron (i.e. a straight stream of electrons will bend).

Mechanism

i. Gas molecules are drawn into a high-vacuum (10^{-5} mm Hg), ionizing chamber.
ii. The cathode bombards molecules with an electron beam.
iii. Molecules lose electrons $\rightarrow$ fragmented into positively-ionized particles.
iv. The ions are accelerated, and subsequently focused into an ion "stream".
v. Then either:
    a. Magnetic field (Magnetic sector mass spectrometer), **OR**
    b. Electrostatic field (Quadropole mass spectrometer)

    These deflect ions, with the resultant mass:charge
vi. There is a spectrum of ratios of mass:charge (dependent on molecules)
vii. A detector to measure mass:charge, and effectively gas concentrations
    a. $\uparrow$ mass:charge = $\downarrow$ deflection
    b. $\downarrow$ mass:charge = ++ deflection

    This is the resultant "mass spectrum".

Interpretation

Molecule	**Detection (mass:charge)**
CO_2	Detected as C^+ charge
N_2O	NO^+, or NO_2^+
Halothane	118
Enflurane	68
Isoflurane	51

The parent compounds, and peaks of the breakdown products are examined.

Advantages	**Disadvantages**
Simultaneous identification of multiple gases (multiplexing)	Big, bulky
Quick ++ (100msec), with 95 % response	Very expensive
Small sample flow rates (1 μlitre, 20 ml/min)	The gas is not returned to the breathing circuit
	Duplicate modules needed for O_2, and CO_2
	"cracking": gas molecules fragment when measured.

Cost and size usually prohibits the presence of more than one of these per theatre complex.

3) Magnetic Susceptibility

Paramagnetic

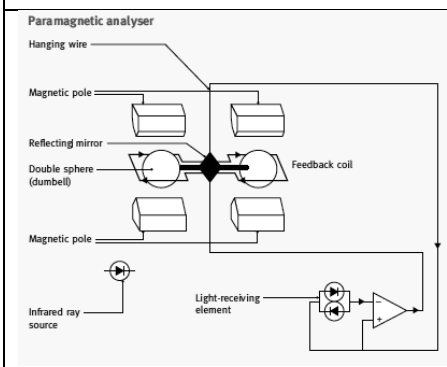
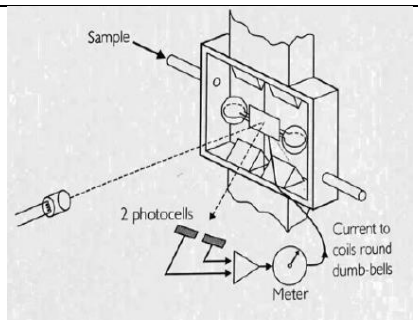
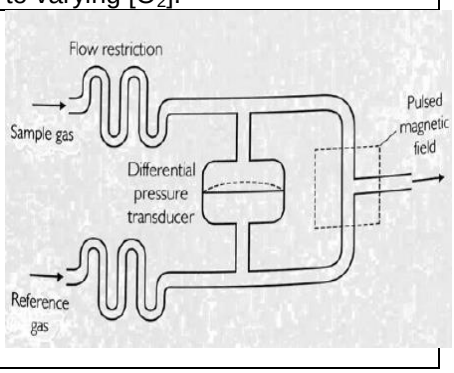
Definition: a substance which is very weakly attracted by the poles of a magnet, but which does not retain any permanent magnetism.

Vs. Diamagnetic (repelled from the magnetic field)

There are three types commonly employed:

1. Deflection type
2. Non-deflection type
3. Pulsed field

Oxygen is paramagnetic (vs. most other gases which are diamagnetic). Oxygen is *strongly* attracted into a magnetic field due to the electrons in the outer valency being **un**paired. Other anaesthetic agents are weakly attracted into the magnetic field.

Deflection Type (older)	**Non-Deflection Type**	**Pulsed Field (Differential pressure transducer)**
∗ 2 glass spheres connected in a dumbbell arrangement ∗ N_2-filled dumbbell is suspended in a filament (to allow for free rotation). N_2 is weakly diamagnetic. ∗ THEN add O_2 which is attracted into the magnetic field (paramagnetic property of O2) ∗ The N_2-filled dumbbell rotates until the force of displacement is balanced by the tension of the filament. The deflection is measured by a mirror. This determines the degree of **light deflectio**n onto a calibrated scale. ∗ The O_2 % is thus determined.	e.g. "Servomax" ∗ 2 glass spheres connected in a dumbbell arrangement (N_2-filled) ∗ Initially in balance (i.e. NO magnetic field) ∗ O_2 sucked into the chamber, changing the magnetic field ∗ O_2 drawn towards the magnetic field zone ∗ Current applied to keep the spheres and dumbbell uniform. The resultant amount of **current** necessary to keep the dumbbells steady gives the indication of the $[O_2]$. The amount of O_2 result in an electrical signal (transduced), and then displayed on a monitor.	e.g. "datex"-type the main advantage is that there are NO moving parts. ∗ Differential Pressure transducer measures the difference of the pressure between two streams of gas. (i.e. a sample, and a reference) ∗ The reference is a known concentration, with known magnetic susceptibility (e.g. O2, air, etc.). ∗ The alternating (on-off) pulsed magnetic field has a frequency of 110 Hz. ∗ Both gases are continuously drawn in across the field. **On:** O_2 molecules are attracted into the field. This drops the pressure at the side of the transducer with ↑ O_2. The differential pressure transducer will move towards the side of > $[O_2]$. **Off:** differential pressure transducer normalizes. The output translates into a voltage change of 20 – 50 μbar, correlating to varying $[O_2]$.
		

4) Chemico-Physical Partitioning

Piezo-Electric Crystals

When an electric potential is applied across a quartz crystal there is a slight contraction in the frequency of the crystal. There are two crystals present in the system.

The first crystal is coated in a Silicon-based lipophilic substance. This encounters a volatile anaesthetic agent, vibrating the crystal, changing the frequency. The vibration is directly related to the solubility of the co-efficient, and determined by Henry's Law: the amount of vapour dissolving is proportional to the partial pressure of the vapour. The degree of vibration determines the amount of the substance. The second crystal is a reference/ standard. Partial pressure is evaluated.

The major problem with this is that the agent cannot be identified (i.e. one is applying knowledge of what has been used), and there is interference from water vapour (which can be eliminated by dissipating the water, by heating).

Sampling systems

	Side-stream	**Main-stream**
Setup	Diverting Delayed response (away from patient) ∗ = 2.5 sec/ 3 m of tubing	Non-diverting No time delay for output Special ETT adapter and module Heat needed to be applied
Volume	150 – 200 ml ∗ Therefore: problem with very low V_T or low fresh gas flow.	Minimal BUT increase by up to 20 ml of V_D
What?	All gas analysers; several gases simultaneously.	CO_2 only
Who?	Intubated and non-intubated patients	Intubated patients
Problems	Catheter clogging, kinking Need water traps, nafion tubing Calibration gas needed	Secretions/ cross-contamination Expensive
Pro's	Disposable components Multiple gases analysed simultaneously	No time delay

There are numerous methods of measurement of gas and vapour analysis – some clinical, and some for research purposes. The important aspect is to ensure that you are au fait with the mechanism of analysis used in the anaesthetic machine that you regularly use (read the manual…), as well as the commonly employed techniques of analysis.

References

1 Recommendations for Standards of Monitoring during Anaesthesia and Recovery. 4th Edition 2007.
2 Goswani A, Cooper R. Measurement of gases (O_2, CO_2, N_2, N_2O and volatile agents). Anaesthesia and Intensive Care Medicine. 2011; 12 (12): 568 – 573.
3 Davis P, Kenny G. Further techniques of gas and vapour analysis. In: Basic Physics and Measurement in Anaesthesia. 2005; Chapter 20: 219 – 226.
4 Garg R, Gupta RC. Analysis of Oxygen, Anaesthesia Agent and Flows in Anaesthesia Machine. Indian Journal of Anaesthesia. 2013;57(5):481-488.
5 Klein D. Gas and vapour analysis. Wits University Part I Anaesthetic Refresher Course. 2008; 103 – 111.
6 Cardoso JF. Gas and vapour analysis. UCT Part I Anaesthetic Refresher Course. 2005.
7 Langton JA, Hutton A. Respiratory gas analysis. Continuing Education in Anaesthesia, Critical Care & Pain. 2009; 9 (1): 19 – 23.
8 Davis P, Kenny G. Carbon dioxide measurement. In: Basic Physics and Measurement in Anaesthesia. 2005; Chapter 19: 211 – 218.

Gas and Vapour Analysis

Technique	O_2	CO_2	N_2O	AxA	N_2	He	Ar	Xe	Advantages	Disadvantages	Problems
Paramagnetic * Pauling * Diff. Pressure * Acoustic	+								Rapid – 150 msec Linear	N_2O effect Water vapour	Temperature-dependent Vibration sensitive Affected by diamagnetic gases Adds N_2 to circuit from reference
Mass Spectroscopy * Magnetic sector * Quadropole	+	+	+	+	+	+	+		Rapid response – 100 msec Small sample = 1 µlitre Multiplexing Quadropole: smaller, variable Mass:charge (ratio)	**Technical:** Bulky, expensive, high vacuum, long warm-up time, high maintenance. Measures only pre-programmed gases. Gas is not returned to the circuit after analysis.	Fragmentation Pressure-sensitive Not mobile Require separate O_2, CO_2 analysis Long response time, but accurate
Infrared Absorption * Photocell * Galoy cell * Luft Detector * Acoustic		+	+	+					Inexpensive, small, rapid response time Gas is returned to the circuit Short warm-up Mainstream for CO_2 Photoacoustic – rapid, better zero, stability	Frequent calibration required Cannot measure O_2 or N_2O	Error in the readings * Overlap spectrum * Water absorption of IR Collision broadening *Frequent* calibration required
Raman Spectroscopy	+	+	+	+	+				All polyatomic gases Multiple anaesthetic agents with good accuracy Returned to circuit	Laser technology	
Piezo-electric				+					Rapid Accurate Return to circuit Short warm-up time	Only has capabilities to measure volatile anaesthetic agents. Cannot differentiate vapours.	H_2O, N_2O: cause interference
Chromatography * Kathorometer * Flame ionization * Ion capture	+	+	+	+	+				Can separate mixtures Able to measure ***very low*** amounts	Non-specific: need fingerprint Non-continuous	Laboratory instrument

Many thanks to Cardoso JF. 2005.[6]

Arterial Transducers and Damping

Dr Gareth Davies

Head of Anaesthesia: Paarl Provincial Hospital
Honorary lecturer- UCT Dept of Anaesthesia & Perioperative Medicine

Introduction

A transducer is a device that converts one form of energy to another form of energy. When measuring intra-arterial blood pressure, the arterial pulse pressure mechanical waveform is transmitted via a column of fluid in the arterial catheter and tubing to a pressure transducer, where it is converted into an electrical signal (hydraulic coupling). This signal is then processed, amplified and converted into a visual display by a microprocessor.

The apparatus required for measuring intra-arterial blood pressure includes the following components:

- Intra-arterial catheter
- Fluid filled tubing
- Pressure Transducer
- Infusion/flushing system
- Signal processor, amplifier and display

Basic Principles

A wave is a disturbance that travels through a medium, transferring energy, but not matter. The simplest waveform is the sine wave.

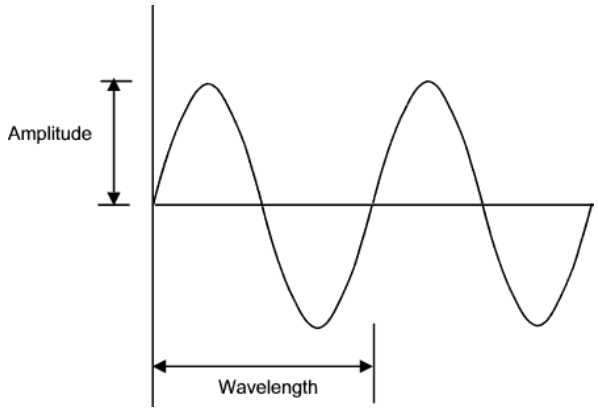


Fig 1: The sine wave

The following key terminology can be used to describe a sine wave:

- *Amplitude*: the maximal displacement from zero
- *Frequency*: the number of cycles per second – expressed as Hertz (Hz)
- *Wavelength*: the distance between two points on the wave that have the same value
- *Phase*: the displacement of one wave in comparison to another

Sine waves are fundamentally important when considering arterial transducers, as any waveform may be represented by combining together sine waves of different frequencies, amplitudes and phases. Thus the complex arterial pulse pressure wave can be broken down into a number of different sine waves (**Fourier analysis**).

This wave consists of a fundamental wave (the pulse rate) and a series of other harmonic waves. These smaller waves have frequencies that are multiples of the fundamental wave (e.g. 2 Hz, 3 Hz, etc.).

A microprocessor performs the function of breaking down the complex waveform into the fundamental wave and at least 10 or more harmonics of higher frequency, to give an accurate representation of the original waveform.

The figures below demonstrate this analysis with two sine waves:

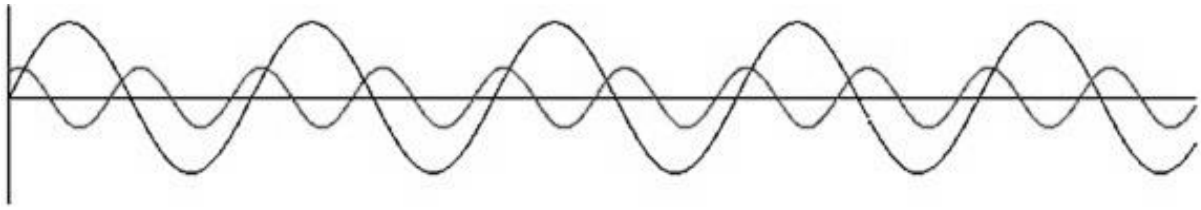


Fig 2: Two sine waves with differing frequency, amplitude and phase

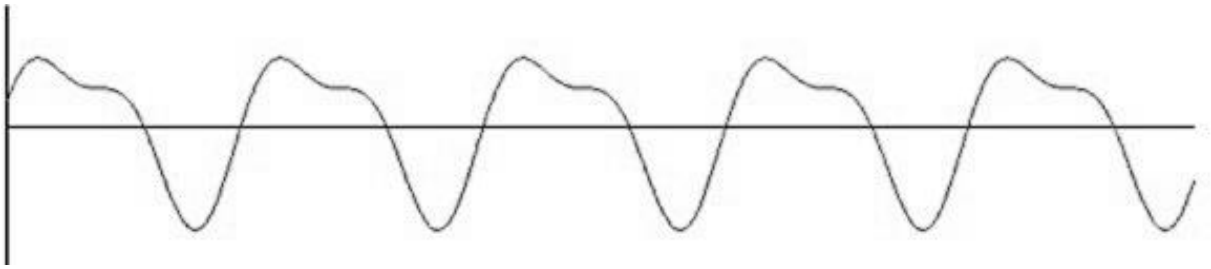


Fig 3: The sum of the two sine waves

Natural frequency and resonance

All materials have a frequency at which they oscillate freely – this is called their *natural frequency*. When a force with a similar frequency to the natural frequency is applied to the system it will begin to oscillate at its maximum amplitude. This phenomenon is known as *resonance*.

This can be demonstrated when pushing someone on a playground swing, where the swing acts as a pendulum. Pushing the swing at its natural frequency makes the swing go higher and higher (maximum amplitude), while trying to push the swing faster or slower produces smaller arcs. This occurs because the energy the swing absorbs is maximised when the pushes match the swing's natural oscillations.

If the natural frequency of an intra-arterial blood pressure measuring system lies close to any of the frequencies of the sine wave components of the arterial pulse pressure waveform, then the system will resonate, producing excessive amplification and signal distortion (amplitude distortion).

The figure below demonstrates that at lower frequencies there is minimal amplitude distortion, however at higher frequencies, where the transduced signal approaches the resonant frequency of the system, significant distortion occurs.

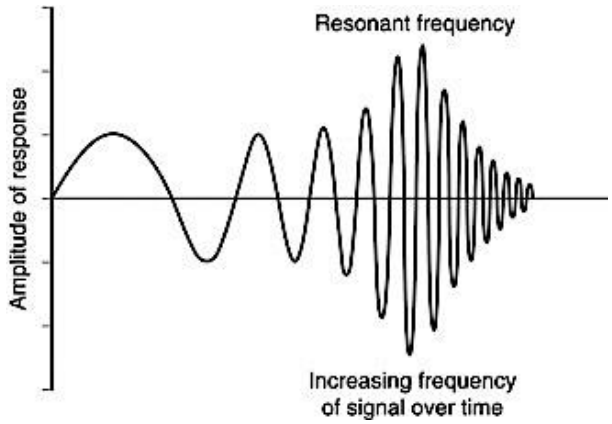


Fig 4: Amplitude distortion at increasing frequency

It is thus imperative that all the components of an intra-arterial monitoring system have a very high natural frequency – at least 8 – 10 times higher than the fundamental frequency of the arterial

waveform (the pulse rate). So for the system to remain accurate for heart rates up to 180bpm, the natural frequency must be at least $\frac{180\text{bpm} \times 10}{60\text{sec}}$ = 30 Hz.
For a fluid filled system the resonant frequency (f_0) can be calculated with the following equation:

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{\pi d2}{4\rho\ell}\frac{\Delta P}{\Delta V}}$$

d = the diameter of the tubing
ρ = the density of the fluid
ℓ = the length of the tubing
$\frac{\Delta P}{\Delta V}$ = the elastance of the system

It is thus clear that when creating an intra-arterial blood pressure monitoring system, that to obtain the highest possible natural frequency, we require shorter, wider, stiffer cannulas and tubing, with lower density fluid.

Damping

In practice the monitoring systems used do not posses a high natural frequency, and *damping* is introduced to decrease or eliminate amplitude distortion. Damping reduces the energy in the system (by creating friction in the fluid pathway) and thus reduces the amplitude of the oscillations. Some degree of damping is required in all systems, but if excessive (overdamping) or insufficient (underdamping), can be a major source of error.

Overdamping can be caused by:
- three way taps
- bubbles and clots
- vasospasm
- narrow, long or compliant tubing
- kinks in the cannula or tubing

Overdamping will result in an under-reading of systolic blood pressure and an over-reading of diastolic blood pressure. The response time of the system is also increased.

In an *underdamped* system pressure waves overshoot, with excessively high systolic blood pressures and low diastolic blood pressures.

Critical damping is the amount of damping required to prevent any overshoot. The damping co-efficient in a critically damped system is 1. However, this results in a system that is relatively slow to respond. A damping co-efficient of 0.64 (termed *optimal damping*) provides the best compromise between responsiveness and distortion.

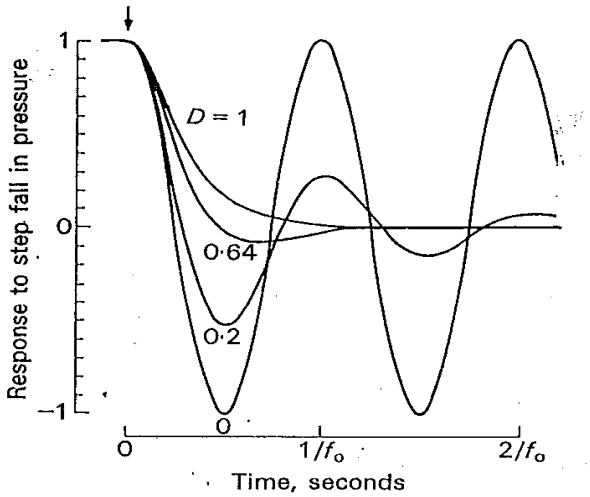


Fig 5: Damping co-efficient (D) – demonstrating the response to a step change in pressure

With a damping co-efficient of 0.64, amplitude is accurately replicated up to 2/3 of the natural frequency (within 2%), and only displays a maximal amplitude distortion of 6% at natural frequency.

In the clinical setting the damping co-efficient of a system can be calculated by using the *"fast flush" test* (see Fig 6 below). The monitoring system is flushed with high-pressure saline via a flush system, creating oscillating waves, resonating at the natural frequency of the system. The trace obtained is then printed and the amplitude of two successive waves measured, allowing the calculation of the amplitude ratio ($^{A2}/_{A1}$). The amplitude ratio is converted to the corresponding damping co-efficient by using available conversion tables. Generally, a lower ratio corresponds to a higher damping co-efficient and a higher ratio represents a lower damping co-efficient.

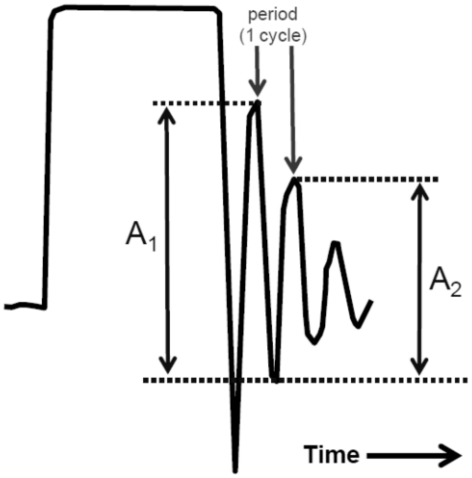


Fig 6: "Fast-flush" test

The fast-flush test is also used to determine the natural frequency of the system. The distance between the peaks of 2 successive waves (one cycle length) is measured and applied to the following formula:

natural frequency = $\frac{paper\ speed\ in\ mm/sec}{length\ of\ one\ cycle\ in\ mm}$

Therefore, in summary: damping is introduced into monitoring systems with low natural frequencies to decrease or eliminate amplitude distortion due to resonance.

Phase shift

As discussed before, the complex arterial pulse pressure waveform can be deconstructed into the fundamental wave with at least 10 additional harmonics. When these waves are summated by a microprocessor, not only their amplitude, but also their phase relationship will affect the displayed waveform. When no damping is present in a system, harmonic waves equal to the natural frequency of the system will be delayed by 90°, whilst those with a very low frequency will demonstrate almost no delay. Waves between these two extremes will demonstrate a variable amount of *phase lag*.

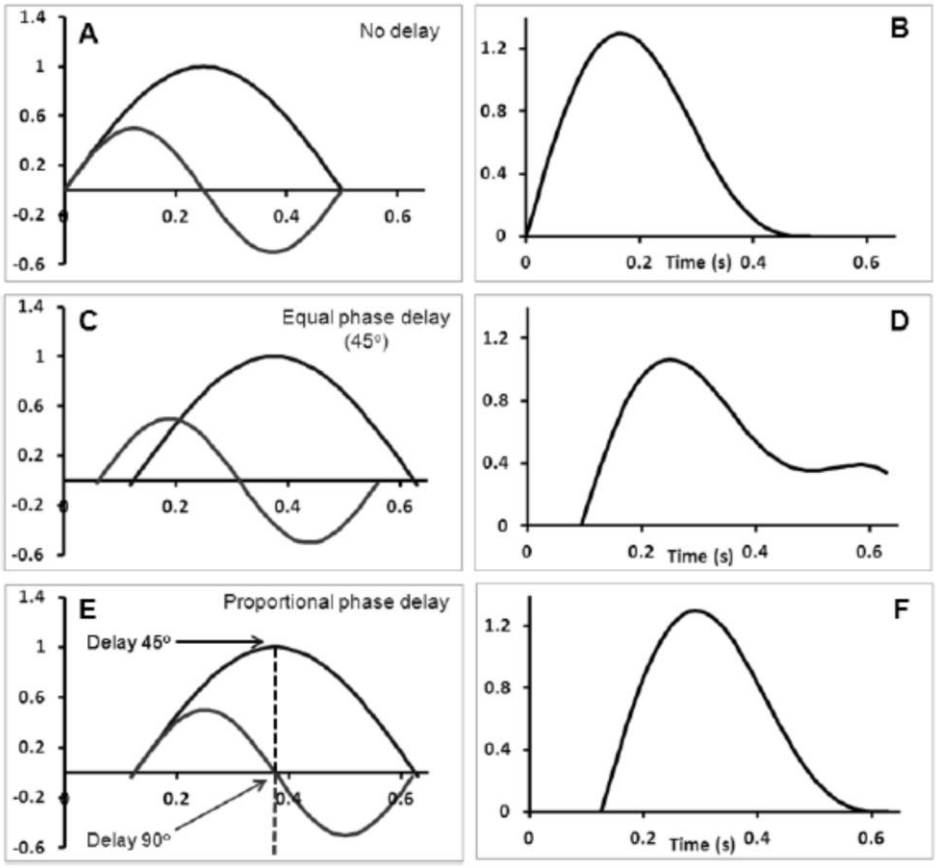


Fig 7: Phase shift – effect of equal versus proportional phase delay

To correct this phase lag, all harmonics need to be delayed in a linear proportion to their frequency. Harmonics with a frequency of 1f (the fundamental frequency) will need to be delayed by 90°, 2f by 180°, 4f by 360°, etc. This produces a delay in the displayed result, but allows the phase relationship of all waves to be preserved

Fortuitously, optimal damping (damping co-efficient = 0.64) provides precisely this proportional delay and allows the accurate summation of all harmonics.

Transducers

In the intra-arterial blood pressure measuring system the arterial pulse pressure is transmitted to a flexible diaphragm by a column of fluid – displacing the diaphragm. The commonest method of measuring this displacement is with a *strain gauge*. Strain gauges utilise the principle that the electrical resistance of wire or silicone increases with stretch. The flexible diaphragm is attached to wire or silicone strain gauges and then incorporated into a *Wheatstone bridge* circuit (see Fig 8 below). With movement of the diaphragm the gauges are stretched or compressed, altering their resistance.

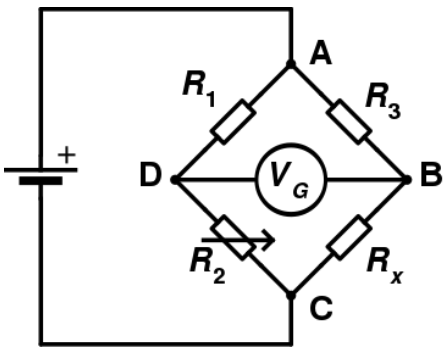


Fig 8: The Wheatstone bridge circuit. When R2/R1 = Rx/R3, the circuit is balanced and no deflection is registered on V_G

These circuits are constructed with 3 resistors of known resistance (R_1, R_2 and R_3) and one of unknown resistance (R_X). The resistance of R_2 is variable and can be adjusted to match the resistance in R_X, until the galvanometer (V_G) produces a zero deflection (voltage between B and D is zero). When pressure is applied to the diaphragm in the measuring system, the two sides of the bridge become unbalanced and current flows. The resistance in R_2 is then altered to maintain a zero balance between B and D, and the magnitude of this change in resistance is proportional to the pressure applied.

Modern Wheatstone bridge circuits have strain gauges in all four positions. The diaphragm is attached in such a way that when pressure is applied to it, the gauges on one side are compressed (reducing their resistance) and those on the other side are stretched (increasing their resistance). The bridge becomes unbalanced and the potential difference generated is proportional to the pressure applied. Using four strain gauges has 2 major benefits: (i) the system is four times more sensitive than a single strain gauge system and (ii) as a change in temperature can affect strain gauges, any temperature change in this system affects all of the strain gauges equally (in a single strain gauge setup temperature change can skew readings).

Zeroing and levelling

Zeroing
Atmospheric pressure must be discounted from the pressure measurement, for the pressure transducer to read accurately. This is achieved by exposing the transducer to atmospheric pressure and then calibrating the pressure reading to zero. As transducers are prone to baseline drift, this should be performed several times a day.

Levelling
The pressure transducer must be set at the level of the patient's heart, to measure blood pressure correctly. By convention this is set at the level of the right atrium, at the 4th intercostal space in the mid-axillary line. If the transducer is placed above or below this level, the hydrostatic pressure exerted by the column of fluid (in this case blood) is being measured in addition to the blood pressure. For a 10cm error in levelling, the measured blood pressure will be over- or under-read by 7.4 mm Hg (over-read if the transducer is too low, under-read if too high).

Summary

The measurement of invasive blood pressure is one of the most valuable clinical tools in the management of critically ill patients. The underlying physical principles may be complex, but are essential to master – allowing the accurate interpretation of displayed values, and the identification of possible sources of error (resonance, damping, phase-shift, etc.)

References

1. Dyer R. The characteristics of the ideal catheter-transducer system for blood pressure measurement. *UCT FCA Part I ARC 2014*
2. Gilbert, M. Principles of pressure transducers, resonance, damping and frequency response. *Anaesthesia and Intensive Care Medicine* 2012 13:1, 1-6
3. Jones A, Pratt O. Physical principles of intra-arterial blood pressure measurement. *Anaesthesia Tutorial of the Week 137* 8 June 2009
4. Joubert, IA. Invasive blood pressure measurement and damping. *UCT FCA Part I ARC 2006*
5. Langton J, Ward M. Blood pressure measurement. *CEACCP* 2007, 7:4, 122-126
6. Morgan Jr GE, Mikhail MS, Murray MJ. Patient monitors. *Clinical Anaesthesiology*
7. Moxham, IM. Physics of Invasive Blood Pressure Monitoring. *SAJAA* 2003, 9:1, 33-38
8. Van West C. Transducers and damping. *UCT FCA Part I ARC 2010*

Bispectral Index
& Other Processed EEG Monitors

Dr Anthony Reed

Head Clinical Department: New Somerset Hospital
UCT Dept of Anaesthesia & Perioperative Medicine

Whilst these notes will focus on the Bispectral Index BIS monitor (Covidien/Medtronic), it is merely the best known, and validated, of a whole family of processed EEG monitors, that include the E-Entropy(GE Healthcare, Helsinki, Finland), the Narcotrend, SEDline (previously Patient State Analyzer), Snap II, and the Cerebral State Monitor.

The essence of BIS ant the other processed EEG monitors is to take a complex signal (the EEG), analyse it, and process the result into a single, dimensionless number that, for BIS and several others, ranges from 0 (equivalent to EEG silence) to 100 (fully awake).

The BIS monitor is based on an initial algorithm developed from just over a 1000 EEG recordings from healthy, adult volunteers exposed to varying levels of sedation and anaesthesia. More recent iterations of the algorithm use input from up to 5000 adults, including patients with stimulation of the CNS caused by surgical stimulation.

The developers then used a multivariate statistical model, including the bispectral (comparing two different waves) and power spectral (comparing the amplitude, or the power, of different frequency waves) variables to identify features that were correlated with the state of anaesthesia, in an effort to develop a monitor that could predict the state of anaesthesia, and /or unconsciousness for clinical use.

Interestingly, BIS values may just reflect a reduced cerebral metabolic rate of glucose, clinically most often produced by hypnotics, but also be seen in other situations that reduce brain metabolism like hypothermia, natural sleep, coma and even unsedated patients under spinal anaesthesia.

So when considering all the processed EEG monitors, one needs to be mindful that they do not measure any effect at the site of anaesthetic action (as we do not know that, and the top layer of the frontal cerebral cortex is unlikely to be the site), and that they do not measure drug concentration. They are at best a correlation of the frontal EEG changes associated with the particular state of anaesthesia, mostly in volunteers, but also with updated algorithms reflecting the EEG changes associated with excitation changes initiated by surgical stimuli.

Background to the processed EEG monitors

History of BIS

In 1996, the U.S. Food and Drug Administration licensed BIS as an aid to monitoring the hypnotic effects of certain anaesthetic agents. In 2003, BIS was licensed " to help guide anaesthetic administration, and may be associated with the reduction of the incidence of awareness with recall in adults during general anaesthesia and sedation."

Effects of anaesthesia on the EEG

The cortical EEG in an awake and conscious patient is typically one with a flat baseline and high frequency, low amplitude waves. With sleep and the administration of most, but not all, anaesthetic drugs; the initial change is increased wave amplitude, followed at larger doses by decreased frequency and increased regularity. So the pathognomonic features are the development of sleep spindles (they look not dissimilar to the Torsade pattern on an ECG) and the entrance of the deep Delta waves. Finally, at very deep levels, periods of isoelectric (flat) EEG interspersed with bursts of undulating EEG activity (burst suppression). However there are some individual differences in the EEG effects of various anaesthetic drugs.

While raw EEG patterns have elements that are anaesthetic-agent specific, concentration- related changes in EEG waveforms are actually quite similar among different agents that potentiate gamma-aminobutyric acid type-A (GABAA) receptors.

Another effect of anaesthetics on the EEG is increased synchronization of the raw EEG waveforms. When the sine waves are consistently in phase with each other, the EEG may be described as synchronized and anaesthetics tend to increase the degree of synchronization in a dose related manner. This is properly called the bicoherence value, and quantifies the extent of phase coupling in a signal. It is also known as bispectral coherency but in EEG processing is termed the “bispectral analysis”.

However a significant portion of the immobility and CNS state produced by anaesthesia is manifest through the drugs’ action on the spinal cord. Cranial EEG analysis may therefore not be a particularly good way to measure the effects of anaesthetic drugs on the spinal cord, and hence is not a particularly reliable method for predicting whether patients will move during surgery.

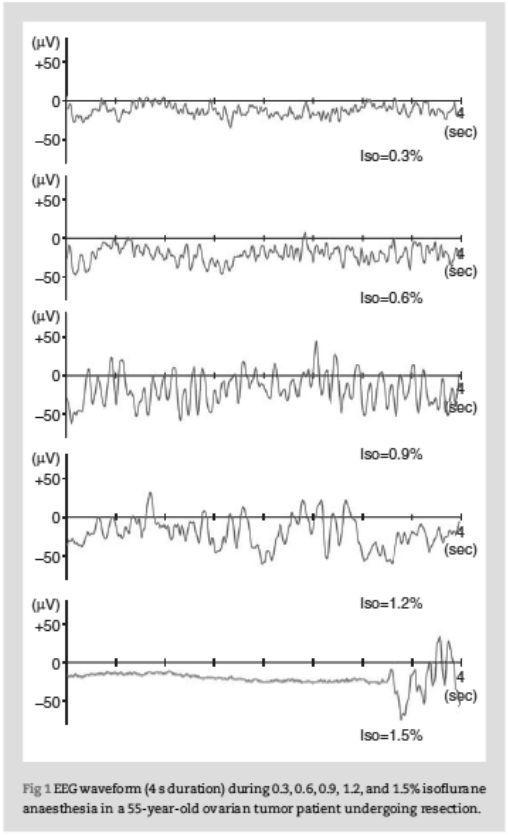
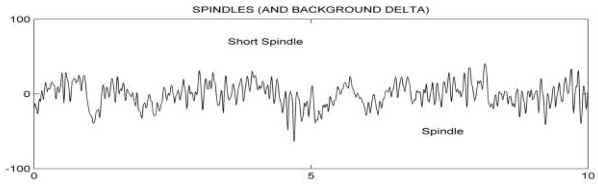


Fig 1 EEG waveform (4 s duration) during 0.3, 0.6, 0.9, 1.2, and 1.5% isoflurane anaesthesia in a 55-year-old ovarian tumor patient undergoing resection.



What we know about BIS

Data acquisition and processing

BIS values are created following advanced analysis of the EEG signal and processing time is an important limitation with computerized EEG interpretation Processing time is substantial, requiring some seconds, although usually less than a minute, and can be easily observed clinically when administering a bolus of an intravenous induction agent to an awake patient. Time delays of 14 to 155 seconds are reported for all of the devices. For a sudden transition from ‘‘general anaesthesia’’ to ‘‘awake,’’ the delays were 15, 30 and 65 seconds for CSM, BIS and Narcotrend, respectively. Recent work published using volunteer anaesthesiologists who received neuromuscular blockade (with no hypnotic agents) with both suxamethonium and rocuronium suggests that the time constant may really be closer to 4 minutes.

Algorithm

The algorithm is a proprietary quantification of several aspects of the EEG, it is empirical and not based upon a physical law or a simple equation.

The BIS value produced is a dimensionless numerical scale that provides a composite correlate depiction (of the overall state, as assessed by the components) of brain activity and is derived from EEG signal processing techniques combining bispectral analysis, power spectral analysis and phase spectral (time domain) analysis, using a proprietary algorithm; including a measure of the degree of sychronisation.

The technique utilises epochs (or frequency ranges – one in the higher range and one in the lower range) that it then assesses the EEG variables, including the spectral edge frequency (the frequency below which x percent of the total power of a given signal is located- typically somewhere between 75 and 95 percent) and median frequency, are calculated solely from the power spectrum. The phase spectrum was traditionally ignored as not being of interest; but the bispectral analysis of BIS is based on power spectrum and phase spectrum and quantifies the coupling of phase angles of different frequencies.

A large component of the assessment appears to be look at this increasing synchronisation of different frequency waves as anaesthesia is increased. Quantifying synchronization requires the analysis of the phase relationships of the component sine waves, and the bispectral analysis including a detailed mathematical technique to analyse the degree of synchronization.

An additional computation is the SynchFastSlow sub-variable, which is derived from the bispectral analysis. SynchFastSlow is defined as the log of the ratio of the sum of all bispectrum peaks in the area from 0.5 to 47 Hz over the sum of the bispectrum in the area 40–47 Hz. BIS is therefore a proprietary combination of SynchFastSlow with a sub-parameter from the frequency domain and a sub-variable from the time domain.

Data presentation by BIS:

- The BIS index range (0–100) is meant to represent a continuum corresponding to the clinical state
- BIS value >97 typically seen in a wide awake patient
- BIS value of 60 has high sensitivity for identifying a state of drug-induced unconsciousness. However it is possible with some combinations of sedatives and analgesics, for an unconscious individual to have BIS value >60.
- BIS value of 40-60 is the recommended range for general anaesthesia, and a value of 55–70 during the last 15 min of surgery
- BIS <40 is a deep hypnotic state
- BIS <30 signifies an increasing amount of EEG suppression
- BIS value of 0 represents an isoelectric EEG signal.

Evidence to support BIS depth of anaesthesia (DoA) monitoring

A meta-analysis of 11 randomized controlled trials of BIS monitoring for ambulatory surgery, comprising 1380 patients, found that the use of BIS monitoring reduced anaesthetic consumption by 19%, reduced the incidence of nausea and vomiting to 32% from 38%, and reduced recovery room stay by 4 minutes.

The four large trials assessing the utility, and efficacy of BIS monitoring in preventing Awareness Associated with General Anaesthesia (AAGA) are the:

- B-AWARE (Myles 2004) trial
    - Multicentre, randomly assigned, adult patients BIS vs Standard anaesthesia (not defined) in high risk patients, showed that BIS decreased awareness in high risk adult patients.

- B-UNAWARE (Avidan 2008) and BAG-RECALL (Avidan 2011) trials
    - Demonstrated that targeted ET AA of 0.7 MAC(control group alert set at 1.3-0.7 MAC) was as good or better than BIS (intervention group alert set at 40-60) in high risk patients
- Mashour (2012) trial
    - Is largest RCT, and was performed in low risk patients unlike the B-AWARE, B-UNAWARE and BAG-RECALL trial
    - Demonstrated that BIS is better than clinical signs in preventing explicit recall, but not better than 0.5 times the age-adjusted MAC in patients under general anaesthesia.

GABAergic relationship

The algorithms have been derived using the common drugs as propofol, midazolam, and isoflurane that are largely GABAergic. Nitrous oxide produces EEG effects that are distinct from the potent inhalational agents, and in most studies nitrous oxide has produced little or no change in BIS or entropy index values. Nitrous is a relatively weak hypnotic, but a good analgesic, and the effects seem similar to that of the opioids, that have relatively little effect on the EEG.

Ketamine is an unusual intravenous anaesthetic because it produces EEG activation, an increase in high frequency activity in the EEG, which often paradoxically increases the BIS index or other EEG-derived indexes, depending upon the dose used. Smaller doses of ketamine may not have a noticeable effect on the BIS index.

Essentially, BIS seems “blind” to agents that are thought to be dominant at non-GABAergic neurones, such as nitrous oxide and ketamine, despite both being established to add to the state of anaesthesia. The addition of ketamine to a standard anaesthetic can even result in an increase in the BIS value, when most would agree that the additional analgesic and hypnotic effects should decrease BIS. Similarly the effect of opiates is also “not seen” by BIS- so the MAC sparing effect of opiates is not detected.

The effects of etomidate and dexmedetomidine on the EEG are not well studied, but the BIS index does appear to track the effects of these drugs.

Challenges with drugs with alternative mechanisms, dementia, opiates, spinal anaesthesia

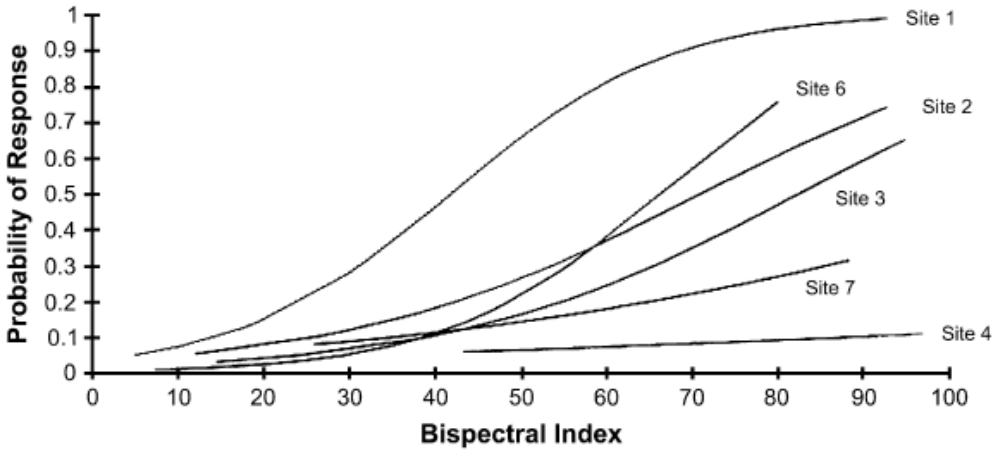


Fig. 1. The probability of movement in response to a surgical stimulus versus BIS index (bispectral index) for seven centers in a multicenter study. Each center employed a different anaesthetic technique, and variable amounts of opioid were used. When isoflurane was the sole anesthetic agent, there was a sigmoid-shaped relationship with a probability of movement of about 50% at a BIS index of 40 (Site 1). For the other six centers in which varying amounts of opioid were combined with isoflurane, propofol, or nitrous oxide, the probability of movement at a BIS index of 40 was reduced to around 10% or less. Site 5 had no movement in response to a surgical stimulus, thus, a relationship to BIS could not be determined. (From Sebel PS, Lang E, Rampil IJ, et al. A multicenter study of bispectral electroencephalogram analysis for monitoring anesthetic effect. Anesth Analg 1997;84:896; (with permission.)

Some clinical situations can significantly influence the BIS value and render it completely unrepresentative. These include:

- Influence of muscle tone (EMG) from the forehead muscles, body motion and sustained eye movements
    - The BIS monitor evaluates the presence of EMG or other high-frequency noise and lights up an EMG signal-strength indicator on the monitor screen alerting the user that the BIS index may be influenced by artifacts
    - Administration of a muscle relaxant (Usually only a small dose is required.) will usually eliminate the EMG activity and restore conditions for reliable EEG signal acquisition. The true BIS index value can then be ascertained.
    - The BIS index value typically falls following administration of a muscle relaxant when an EMG artifact is present, since the EMG activity usually elevates the BIS index value.
- Potential artifacts may be caused by
    - Poor skin contact (high impedance), muscle activity or rigidity
    - Head and body motion, sustained eye movements,
    - Improper sensor placement
    - Unusual or excessive electrical or mechanical interference caused through
        - Electrocautery
        - Cardiac atrial pacing
        - Vibrations from forced air warmers
- BIS values should be interpreted cautiously with certain anaesthetic combinations, such as those relying primarily on either ketamine or nitrous oxide/narcotics to produce unconsciousness.
- In some instances high doses of opioids can cause a paradoxical increase in the processed EEG, while low doses of ketamine can result in increased high frequency EEG activity.
- Because the BIS algorhythm was developed using healthy volunteers with normal EEG patterns, any pre-existing neurologic disorder that exhibits abnormal EEG waveforms can affect the BIS.
    - Patients with Alzheimer's or vascular dementia can show an increase in the slow wave activity of the EEG, associated with a lower mean awake BIS.
    - Patients with cerebral vascular disease may have cerebral ischaemia leading to cortical inactivation leading to EEG slowing or a decrease in the BIS
- Some serious clinical conditions such as profound hypotension and severe hypothermia (hypothermia reduces the BIS by approximately one BIS unit per degree Celsius)

Cost effectiveness

Cost saving could come from two sources- less drugs, shorter theatre and recovery time; and litigation savings if risk of awareness is decreased. One study showed that the costs savings in drugs and recovery time was the equivalent of 0.18 Euro/min, and the costs of the BIS consumables was 14.01 Euros.

Other studies have shown similar outcomes- costs are increased with BIS monitoring and the potential savings do not overcome the additional cost. Currently it cannot be said that BIS is cost-effective.

Concerns

MAC for anaesthetic agents is a measure of the ED_{50} for awake, movement or BAR and is a measure of equivalence for all agents in terms of the desired clinical effect. This has been described as "normalising" the population response between different inhalational agents. So for MAC there is shared probability of 50% of patients not responding/moving in response to the specified stimulus. If a DoA monitor shared this accuracy then for a given value for all anaesthetic agents, it should match the MAC value. This does not hold true so that the BIS value is different for MAC with each inhalational agent, varying as much as BIS =35 for some agents and BIS=60 for others.

The differences between inhalational agents and TIVA with propofol are even more stark, with the probability of unconsciousness at a BIS value of 70 being 50% with isoflurane and only 15% with propofol.

Because healthy adult EEG data were used to authenticate the BIS algorithm, it cannot automatically be extrapolated to young children, as the paediatric EEG only approaches the adult pattern by about 5 years of age. However, it does appear that BIS may be valid in children older than 1 year of age.

Current status of processed EEG DoA monitoring

Product monologue from manufacturers states:

> "Reliance on BIS values alone for intraoperative anesthetic management is not recommended"

The UK National Institute of Heath Research Health Technology Assessment performed a Technology Assessment Report (2012) on EEG based monitors and concluded that:

> "The available evidence on the impact of the technologies on reducing the likelihood of intraoperative awareness is limited. Overall, [EEG-base monitors are] not associated with a statistically significant reduction in intra-operative awareness in patients classified as at higher risk."

Current Opinion in Anesthesiology (2016)

> "Current research suggests that processed EEG monitors may be most useful in specific patient populations, such as TIVAs and in patients with hemodynamic compromise that requires the clinician to minimize the concentration of the vaporized agents. Alternatively, measuring ETAC and maintaining it greater than 0.7 age-adjusted MAC can prevent awareness while being most cost-conscious."

A 2014 Cochrane review found that:

> "Four studies in 7761 patients, that used clinical signs as a guide to anaesthetic administration in standard practice, as the control group, demonstrated a significant reduction in the risk of awareness with BIS monitoring.
>
> Four studies with a total of 26,530 patients, compared BIS monitoring with end tidal anaesthetic gas (ETAG) monitoring as a guide to management of anaesthesia and they did not demonstrate any difference in terms of intraoperative awareness"

There is a movement developing that suggests that rather than provide processed EEG data, that anaesthetists should rather have knowledge of changes in the raw EEG during anaesthesia, and that this it could better help them judge the adequacy of EEG indices and enable them to respond more rapidly and confidently in circumstances where equipment algorhythms provide misleading indications. There is developing evidence that after just one or two weeks training, it seems likely that most anaesthetists should be able to reliably judge and interpret the effects, and the state of anaesthesia on the raw EEG.

Recommended reading:

1. Editorial. Was NAP5 "NICE" enough; where next for depth of anaesthesia? Anaesthesia 2015; 70:511-527.
2. Review. Advances in awareness monitoring technologies. Current Opinion in Anesthesiology 2016; 29: 711-716.
3. Review. Depth of Anaesthesia Monitoring. Anesthesiology Clinics 2006; 24: 793-822.
4. Editorial. Mind the gap: Attitudes towards intraoperative brain monitoring. Anaesthesia and Analgesia 2014; 119:1022-25.
5. BAG-RECALL trial. Estimation of the Bispectral Index by Anesthesiologists. Anesthesiology 2011; 114:1092-1101.
6. Response of bispectral index to neuromuscular block in awake volunteers. British Journal of Anaesthesia 2015; 115:95-103.
7. National Institute for Clinical Excellence guidance on measuring depth of anaesthesia: limitations of EEG-based technology. British Journal of Anaesthesia 2013; 110: 325-8.
8. How electroencephalography serves the anesthesiologist. Clinical EEG and Neuroscience 2014; 1-11.
9. Prevention of Intraoperative Awareness with Explicit Recall in an Unselected Surgical Population. Anesthesiology 2012; 117:717-725
10. B-UNAWARE trial. Anaesthesia awareness and the bispectral index. NEJM 2008; 358: 1097-1108.
11. Bispectral index monitoring to prevent awareness during anaesthesia: The B-Aware randomized controlled trial. Lancet 2004; 363: 1757-63.

Defibrillators

Dr Marcin Nejthardt

UCT Dept of Anaesthesia & Perioperative Medicine

Defibrillation is a process of passing a *current* through the myocardium to depolarize a critical mass of heart tissue facilitating an arrest of electrical activity and relying on its automaticity to resume orderly electrical activity.

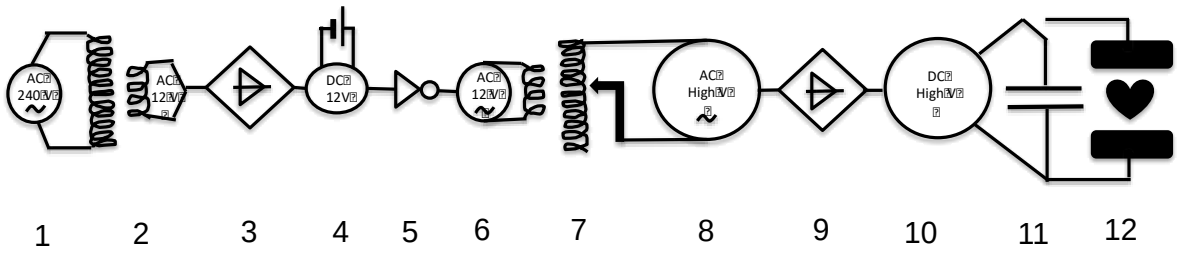
In essence, through some fancy electronics based on fundamental physical principles defibrillators convert a low voltage, low current input (battery/mains) and produce a short duration, high intensity current output to facilitate defibrillation.

Defibrillator components:

1. Power source
    a. Mains
    b. Battery
2. Capacitor
3. Inductor
4. Switches
5. Paddles

1. Power Supply

Schematic of power supply to charge the defibrillator paddles



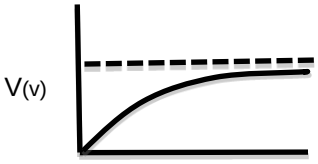
1. Mains – 240V AC 50/60Hz
2. Step down transformer – 240V AC to 12V AC
3. Rectifier – converts AC to DC by utilitizing 4 diodes in an H-bridge formation
4. Rechargeable 12V Lead Acid Battery (charged by constant current with constant voltage)
5. Inverter – converts DC to AC. Like the rectifier it relies on 4 diodes in H-bridge formation to regulate direction of flow. High frequency electrical switches turn on and off to allow pulsation of current to alternate in a back and forth direction.
6. AC 12V
7. Variable step up transformer – allows us to set the desired energy (Joules)
8. Generation of high voltage AC
9. Rectifier
10. Generation of high voltage DC
11. High voltage capacitor
12. Paddles

The defibrillator can thus be charged by one of two power sources, from the battery (12V DC) or the mains (240V AC). Unfortunately the voltage from both sources is too low to generate an adequately high potential difference across the capacitor for effective defibrillation. To correct this, the circuit relies on rectifiers, inverters and transformers to achieve the desired voltage.

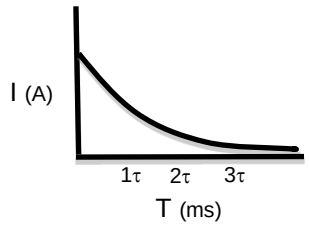
2. Capacitors

A battery's primary function is to store energy and not generate current. The current that can be generated even with a low resistance conductor is far too low to provide an effective current for defibrillation. It is for this reason that capacitors are incorporated. They are essentially a pair of conducting plates separated by an insulator known as a dielectric. The capacitor's function in the defibrillator is to store charge (Q) and then discharge it in a short burst of high current.

Graphs of *Voltage* and *Current* vs *Time* during the charging of a capacitor



Voltage across a capacitor tends towards the battery voltage (dashed line).



$I = \frac{V_0}{R} e^{-t/RC}$ represents the current in the circuit.

τ is the *time constant* for charging and discharging the capacitor.

For any capacitor $\tau = RC$

Where:
R = resistance
C = capacitance

How *fast* the capacitor charges depends on the *time constant* of the circuit. Since RC is the time constant of the circuit with the capacitor anything that decreases the product will *decrease* the charge time

1. Capacitor size (small capacitance faster charge)
2. Low resistance (fast charge)

An indication of how much charge a capacitor can hold is determined by its *capacitance* (units Farad, but usually in the clinical context μF are used)

$$C = k\, \varepsilon_0 \frac{A}{d} \qquad (1)$$

Where:
K= dielectric constant
ε = permittivity of free space (constant)
A = area
D = distance

To increase the storage of a capacitor manufacturers manipulate the above equation.

1. Increasing dielectric constant - by inserting a polar insulator like glass or plastic which during charging aligns its molecular orientation to attract more electrons onto one capacitor plate and repel electrons from the other plate.
2. Increasing the surface area of the plates
3. Decreasing the distance between the plates

External defibrillator capacitors have a capacitance range of 32-500μF. In addition, the capacitor must also be HV (high voltage) with a typical DC voltage range of 800-6000V to accommodate the high potential difference needed for adequate current flow.

$$C = \frac{Q}{V} \qquad (2)$$

Where:
C = capacitance
Q = charge
V = voltage

Thus a capacitor rated as high voltage has the ability to store more charge ($Q = CV$).

A volt is the *energy* required to move 1 coulomb of charge against a magnetic field, thus rearranging the equation we get

$W = VQ$ (3)

Where:
W = work (which is energy measured in Joules)
V = voltage
Q = charge

Work is the energy we set for a defibrillator to deliver. It represents the work done in moving a charge through a potential difference. Because of resistance in the system, the actual workable energy released to the patient is described by:

$$E = \frac{1}{2}QV. \quad (4)$$

Combining equation (2) and (4) from above, the amount of energy delivered to the heart expressed in terms of capacitance and voltage is:

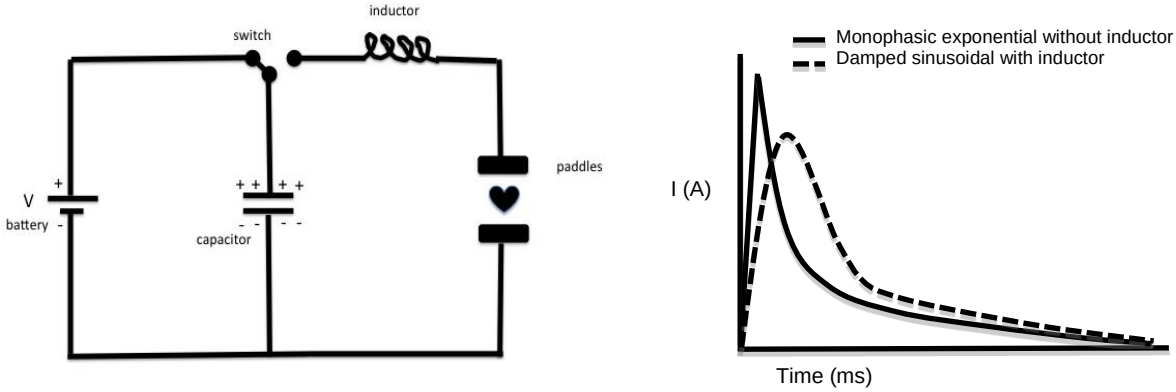
$$E = \frac{1}{2}Cv^2. \quad (5)$$

Where:
E = energy delivered (in Joules)
C = capacitance (constant for a capacitor as per equation (1))
V = voltage

When we dial up the different Joule settings for the defibrillator to deliver, we are effectively altering the potential across the capacitor using a variable step–up transformer.

Discharging a defibrillator and associated waveforms

Monophasic defibrillator



Current flow by convention is from the positive terminal of the battery (long thin bar) to the negative terminal (short thick line). In reality however it is the flow of electrons in the opposite direction that is actually responsible for current.

Charging the capacitor in the above diagram occurs in a clockwise direction establishing a positive charge on the upper plate and a negative charge on the lower plate. In discharging the capacitor after switching from the battery to the paddle circuit, the current flow is in a clockwise direction.

The inductor in the monophasic defibrillator prolongs current flow across the paddles for an appropriate length of time to defibrillate the heart. Additionally the inductor lowers the peak current resulting in less injury to the myocardium.

3. Inductors

An inductor is essentially a coil of wire that acts as an electromagnet with the ability to temporarily store energy in the magnetic field it generates when a current flows through it.

What's fascinating about a coil is that magnetic fields from different parts of the wire intersect. With a change in current (increase or decrease) there is an apparent movement of the magnetic field in relation to the coil which induces an electromagnetic force (emf) in the opposite direction to the current. This is known as *inductance (L)*. By varying the source voltage we can cause a change in

current to induce an opposing electromagnetic force (emf) in the coil which can then be used as a temporary store of electromagnetic energy.

The mathematical build up of this potential energy is identical to that of a capacitor taking approximately 3 time constants to reach a plateau. The time constant for an inductor is calculated by $\tau = \frac{L}{R}$.

On reaching the plateau phase, the emf falls to zero which corresponds to no change in current flow (gradient of the *current* vs *time* curve is zero (horizontal flat line)) and the inductor offers almost no resistance to current and acts as a conductor in series.

By placing a resistor (load) in the circuit with an inductor, we can generate a voltage across the resistor which will depend on the size of that resistance (V=IR) i.e. inserting a high resistance will generate a high voltage.

$$L = \frac{\mu_0 N^2 A}{l}$$

Where:
L = inductance (unit Henries)
μ_0 = magnetic constant
N = number of turns
A = cross sectional area of coil
l = length of coil

Manufacturers can increase the inductance of the coil by wrapping it round a magnetic core thereby increasing the magnetic constant (μ_o). Increasing the number of turns has an exponential effect (N^2) and having a coil with large turns (increased area) increases inductance. By keeping the length of the coil to a short length also increases inductance.

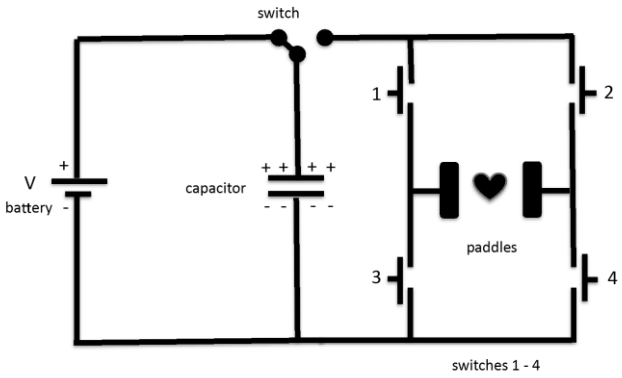
Energy stored by an inductor is: $E = \frac{1}{2} LI^2$.

Where:
L = inductance
I = current

Therefore the greater the inductance of the inductor the greater the potential energy it can store.

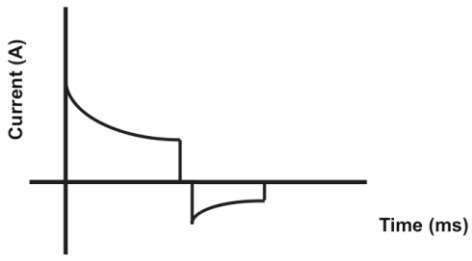
Biphasic defibrillators

The advantage of these machines compared to the monophasic defibrillators is that the peak current delivered is even lower (less myocardial damage) without compromising on the time to first effective defibrillation. A need for lower current equates to a smaller capacitor and hence less bulky defibrillator at a better price. The use of switches arranged in an H bridge allows for bidirectional current flow.



Switches 1 – 4 form an H bridge which allows a voltage to be applied in either direction. It is this switch configuration that allows the defibrillator to discharge current from left to right and then reverse direction from right to left.

Biphasic Truncated Exponential Waveform



Biphasic = two direction (above/below zero line)
Truncated = current flow is cut-off
Exponential = rate of current decline is non-linear

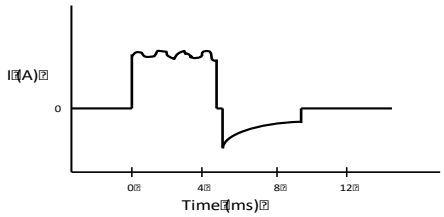
The capacitor is discharged by flipping the switch from the battery to the paddle circuit and turning on switches 1 and 4. After a period of time (determined by the resistance of the patient) switches 1 and 4

are turned off and the current abruptly ceases to flow (first truncation). This truncation allows for enough charge to be left on the capacitor to facilitate additional current flow in the opposite direction. A brief period of no flow is followed by a period of reverse flow initiated by turning on switches 2 and 3. The second truncation occurs before the capacitor is fully discharged. Because a low current has the ability to induce fibrillation, the second truncation is initiated at a predetermined current threshold level by turning off switch 2 and turning on switch 1. This effectively short-circuits (bypasses the paddles) through switch 1 and 3 and internally discharges the remainder of the charge on the capacitor.

The effect of patient resistance on current flow

A 200J setting on a defibrillator has the effect of delivering slightly differing currents for patients of differing resistance (Ohm's Law). Thus a thin patient, with a low tissue resistance will experience a *higher* current flow for a *shorter* period compared to a large patient with much adipose tissue who will experience a *lower* current for *longer*. Important to note is that the area under the *Current (I)* vs *time (ms)* graph represents Q, the charge on the capacitor ($Q = It$) and thus both patients (despite different resistance) will receive the same predetermined charge. Biphasic truncated exponential defibrillators can adjust and deliver an appropriate current to patients with an impedance range of 25 – 180Ω.

Basic rectilinear waveform



In these machines the period of defibrillation is kept constant at 10s. The machine attempts to maintain a constant current delivery during the first phase by adjusting its own internal impedance dependent on the patient's resistance. Based on Ohm's Law for a patient with a low resistance additional internal resistance will be added to keep current flow constant.

4. Paddles

Up to now we have focused on how the machine is able to setup the energy required to deliver the optimal current across the heart. Crucially, paddle placement needs to be correct (apex to right of sternum below clavicle or in an anterior to posterior orientation) to ensure that the current from one paddle passes through the heart. Appropriate downward pressure (±100N (10kg)) is also important. The paddles also need to be the correct size to ensure appropriate current density (adults 10-13cm, children 8cm and infants 4.5cm).

$$\textit{Current (I) density} = \frac{I}{area\ of\ paddle}$$

Too small an area of contact will lead to burning of skin and tissue. Contact between the paddles and skin is coupled through the use of a gel to decrease impedance. This allows for adequate current delivery to the deeper structure (heart) and avoids burning of superficial structures (skin).

Conclusion

Understanding the physics of a defibrillator is rooted in having a clear appreciation of electrical concepts. Delegates are encouraged to complement this reading by exploring the basics of electricity particularly on the differences between alternating and direct current and the workings of transformers. Additionally, basic mathematical concepts of exponential processes and time constants are mentioned, but more in depth understanding (which was beyond the scope of this lecture) is encouraged. Finally, capacitors, inverters and rectifiers are introduced in developing an understanding the function of a defibrillator.

References

1. S.K. Singh, R. Ingham and J.P. Golding. Basics of electricity for anaesthetists. CEACCP vol 11. Number 6. 2011.
2. D.J. Williams, F.J. McGill and H.M. Jones. Physical principles of defibrillators. anaesthesiauk.com
3. Philips Heartstart XL service manual
4. Howequipmentworks.com/defibrillators
5. Youtube Physics Videos by Eugene Khutoryansky

Temporary Pacemakers

Dr Owen Porrill

UCT Dept of Anaesthesia & Perioperative Medicine

Temporary pacemakers are usually inserted as part of the urgent management of severe arrhythmias, usually bradyarrhythmias. They may also be inserted for prophylactic indications e.g. patients with a high-grade AV block undergoing surgery. Rapid atrial pacing is also used to decrease cardiac output e.g. during TAVI placement. The sites of insertion we are most likely to encounter as anaesthetists are transvenous or epicardial after cardiac surgery. Other routes of pacing are transcutaneous or oesophageal, which require different equipment and will not be discussed in detail in this course.

Temporary pacemakers consist of a pulse generator (box) and insulated wires, also known as leads. The box contains the controlling circuitry and a battery and has two basic functions. The first is to sense electrical activity in one or more cardiac chambers. The other is to pace the atrium, ventricle or both.

Leads

Leads may be transvenous/endocardial or epicardial/myocardial. Endocardial leads are introduced via the venous system and are placed in contact with the endocardium. Epicardial leads are applied to the external surface of the heart and are introduced through the chest wall during cardiac surgery (atrial on the right, ventricular on the left). Epicardial leads are usually sutured in place.

Transvenous leads may be flexible, semifloating (with or without a balloon tip) or rigid. They may also be incorporated into a pulmonary artery catheter. Rigid wires are inserted under fluoroscopic guidance, the others are inserted blindly or under ECG guidance. The best insertion sites are the left subclavian or right internal jugular veins, due to the curvature of the lead. The femoral vein may occasionally be chosen in anticoagulated patients. The side of the dominant hand is often used as a permanent pacemaker may subsequently be placed on the non-dominant side. Ultrasound may be very useful in placement of the leads, both for vascular access and to visualise the lead in the cardiac chambers and to see ventricular contraction in response to stimulation. Ultrasound may also be useful in diagnosing complications of the procedure e.g. haemopericardium or pneumothorax.

Fixation of endocardial leads may be passive or active. Passive fixation occurs by means of tines (small prongs) which become enmeshed in the trabeculae of the ventricular apex or the pectinate muscles of the atrium. Active fixation occurs by means of a helix which is screwed into the myocardium.

The pulse generator, leads and body tissue are all components of the electrical circuit. Most leads nowadays are bipolar, meaning that the positive and negative electrodes are contained within the lead. The current flows from the pulse generator to the tip of the pacing lead (cathode). The anode is a metallic ring a few millimetres from the tip. In unipolar leads the current flows from the pulse generator to the cathode and back to the box that acts as the anode. Unipolar leads require a higher current due to the increased impedance. Unipolar leads are also more sensitive to electromagnetic interference and have a higher risk of diaphragmatic or chest wall muscle pacing. Unipolar leads usually cause a larger pacing spike on the surface ECG.

Leads may be inserted into the atrium, ventricle or both, depending on the indication for pacing.

External Pacing Generator
This will have the following controls and capabilities:

1) Rate: Controls heart rate
2) Output: This controls the output current in milliamperes, usually between 0,1 and 20 mA
3) Sensitivity: Senses native electrical activity in the cardiac chamber. The sensitivity control sets a threshold at which pacing will be inhibited.

Rate

In practice, the rate is usually set at 80-90 per minute. The rate is sometimes reduced to 40 before pacemaker removal as a backup rate.

Sensing

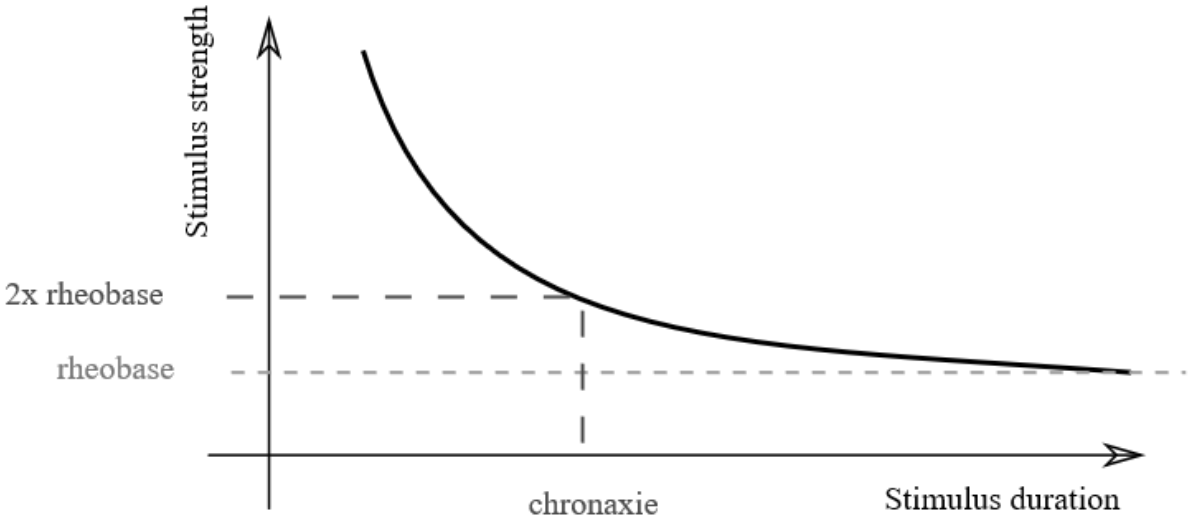
The pacemaker senses intrinsic cardiac activity. This is measured in millivolts. The sensing threshold can be adjusted by the user. Remember that increasing sensitivity means decreasing the threshold in mV and vice versa (detection of smaller potential difference equals higher sensitivity). The sensing threshold must be checked if the pacemaker is to be used in demand or synchronous mode. The pacemaker should be in VVI, AAI or DDD mode. Set the pacing rate 10 beats per minute less than the intrinsic rate and the output at 0,1 mA. Set the sensitivity at its highest (demand mode). The pacemaker should only sense and not pace. Decrease the sensitivity until the pacer starts pacing and the sensing indicator stops flashing. Increase the sensitivity until the sense indicator starts flashing and the pace indicator stops flashing. This is the pacing threshold. Set the sensitivity at half of the threshold value for a margin of safety, but to avoid undersensing. If there is no underlying rhythm, set the sensitivity at 2 mV.

Sensing filters exclude myopotentials, T waves, pacemaker activity on a different channel (crosstalk) and other extraneous signals (oversensing).

The asynchronous mode is potentially dangerous and should only be used in a highly monitored environment. We often use the asynchronous mode in theatre to avoid oversensing due to electromagnetic interference. If the pacemaker delivers a stimulus during ventricular repolarisation (R-on-T phenomenon), ventricular fibrillation may be precipitated. The pacemaker should always be reprogrammed to a synchronous mode before the patient moves to a general ward.

Physiology of pacing

An electrical stimulus will cause depolarisation and subsequent action potential propagation if the energy delivered exceeds a certain threshold. The energy delivered is a function of current delivered and the time over which the current flows. Each excitable tissue has its own threshold.
The energy required is described by the strength duration curve. Electrical capture will occur in the area to the right of the curve. The rheobase is the minimum current required to elicit a response if the stimulus duration is very long. The chronaxie is the time required to elicit a response at a current twice the rheobase. These are not terms you will use on a day-to-day basis, but they are important to electrophysiologists and engineers designing equipment such as pacemakers and nerve stimulators. Practically, a pacemaker generates a continuous current for about 1,5 milliseconds, which approximates the chronaxie of cardiac muscle. The time is not adjustable, but the user can control the current output.



To set the appropriate output, you need to determine the pacing threshold. It is unsafe to check the threshold in the absence of an underlying rhythm as you may not be able to re-establish capture. If it is safe to check the threshold, set the pacing rate at 60-80 beats per minute, or at least 10 beats per minute faster than the patient's intrinsic rate. Start with a high current output and decrease the current until capture is lost. Increase the current until consistent capture is demonstrated (each pacing spike is followed by a p wave or QRS complex). This is the threshold value. The output is then set to 2-3 times this value to ensure reliable function. If it is >10 mA, a smaller margin of safety is used to avoid accelerating myocardial fibrosis. The threshold should be checked regularly as it can rise over time as an inflammatory reaction develops around the lead implantation site. Temporary endocardial wires may be steroid eluting to decrease the inflammatory response and are generally reliable for periods of up to two weeks. Epicardial pacing wires have a shorter lifespan, often failing to pace and/or sense by day 5. Unnecessarily high pacing output will decrease the life span of the temporary lead and increase the inflammatory response in the myocardium.

To reassess the underlying rhythm, it is important to turn down the pacing rate and allow the native rhythm to manifest. If the current is decreased to allow the underlying rhythm to emerge, the threshold for recapture may be dramatically increased and it may be impossible to resume pacing (Wodinsky effect). The underlying rhythm should always be assessed and must be more than 50 per minute and the patient must remain haemodynamically stable before sensitivity and capture thresholds are assessed (in that order).

Pacemaker modes

The NBG code for pacemakers applies to temporary pacemakers. The first 3 positions are the most relevant, but biatrial or biventricular pacing is occasionally used after cardiac surgery.

The revised NASPE/BPEG generic code for antibradycardia pacing

I	II	III	IV	V
Chamber(s) paced	Chamber(s) sensed	Response to sensing	Rate modulation	Multisite pacing
O = None	O = None	O = None	O = None	O = None
A = Atrium	A = Atrium	T = Triggered	R = Rate modulation	A = Atrium
V = Ventricle	V = Ventricle	I = Inhibited		V = Ventricle
D = Dual (A+V)	D = Dual (A+V)	D = Dual (T+I)		D = Dual (A+V)

Common modes used are AAI, VVI, DDD, DOO and AOO, depending on the scenario. The asynchronous modes (AOO, DOO, VOO) are generally only used in theatre when diathermy is used or during an emergency when there is malfunction of one of the more advanced modes. Atrial pacing is used for bradycardias with intact ventricular conduction. DDD is the most useful mode as it preserves AV synchrony. VVI has the advantage of simplicity, requiring only 1 lead to be inserted. VDD is occasionally used in cases of AV nodal block with an intact sinus node. Triggered modes (VAT, AAT, DAT) are not commonly used in temporary pacing.

Atrial overdrive pacing may be used for supraventricular tachycardias, but that will not be discussed in detail here.

Typical settings for postoperative patients
Rate: 80-90 bpm
Output: 10mA (atrium) 10-20 mA (ventricle)
Sensitivity: 5 mV (atrial 0,4-10 mV, ventricle 0,8-20 mV)
AV delay: 150 ms (or auto depending on rate). May be changed to optimise cardiac output
PVARP: 250 ms (or auto depending on rate)

Complications

Complications of temporary pacing include complications of central venous cannulation as well as complications related to the pacemaker. These include mechanical complications such as perforation of the heart or ventricular septum, infection, looping, damage to the tricuspid valve, misplacement or

interference with vascular anastomoses. There may also be problems of functionality e.g. failure to sense or capture.

Failure to capture is recognised when pacing spikes are seen, with no cardiac activity. This is usually caused by lead failure, but other causes may be hypoxia, ischaemia, acidosis, alkalosis, electrolyte abnormalities (particularly hyperkalaemia) and antiarrhythmic drugs. A temporary solution in an emergency is to reverse the polarity of the electrodes. If a bipolar lead is used, the anode can be used as the cathode and a return electrode inserted subcutaneously.

Failure to sense has similar causes to failure to capture.

Crosstalk can occur with dual chamber pacemakers, with pacing of one chamber inhibiting the pacing of the other chamber. This most commonly manifests as atrial pacing misinterpreted as ventricular activity, resulting in ventricular standstill. This can usually be addressed by increasing the ventricular sensitivity and/or decreasing the atrial output. Modern pacemakers have two safety features to prevent ventricular standstill. Ventricular blanking ensures that ventricular activity is ignored for a short period after atrial depolarisation. If the pacemaker detects ventricular activity after the ventricular blanking period but before it would expect with normal AV conduction (i.e. during the cross-talk sensing window), it assumes there is either cross-talk or a PVC. It then emits a ventricular spike slightly early after the next atrial depolarisation (i.e. it decreases the AV delay of the next cycle) to avoid ventricular standstill. This is known as ventricular safety pacing.

Pacemaker mediated tachycardia can occur in DDD or VDD modes and has various causes. It most commonly occurs when there is atrial sensing of a ventricular spike. This is interpreted as atrial depolarisation and triggers ventricular pacing and subsequent tachycardia. The pacemaker employs an atrial blanking period to prevent this, during which it will not sense any ventricular depolarisation. It may also occur due to retrograde p waves via an accessory pathway setting up a re-entrant circuit. PVARP (post ventricular atrial refractory period) can be manually prolonged on the pacemaker in this scenario to prevent ventricular pacing in response to the retrograde p wave. Changing to VVI or DVI mode temporarily should also terminate the tachycardia, but this should only be employed temporarily while a definitive solution is found.

Tachycardia can also occur due to inappropriate tracking of a high atrial rate. Modern pacemakers will usually set the upper rate automatically. When the rate reaches this limit, the pacemaker introduces an artificial Wenkebach phenomenon, increasing the AV delay until the atrial spike falls within the PVARP and the impulse is ignored. When this occurs, the tachycardia should be treated or the mode changed to DDI.

Microshock is a danger as these leads are connected directly to the heart. All power sources should be ungrounded and any sources of static should be removed (including Get Well Soon balloons!). Gloves should be used when handling the wires. Anyone handling the wires should touch a large metal object (e.g. the bed) to discharge static build-up before handling the wires.

Removal of the leads can lead to haemorrhage and arrhythmias, thus patients should be appropriately monitored during and after removal.

Endoscopic Airway Equipment

Dr Ross Hofmeyr

Inaugural UCT-Storz African Fellow in Airway & Thoracic Anaesthesia
UCT Dept of Anaesthesia & Perioperative Medicine

Contents

Introduction

Safe, reliable airway management is integral to the practice of anaesthesia. While some procedures may be accomplished without advanced instrumentation of the airway, there are many in which the ability to accurately position endotracheal and endobronchial tubes, bronchial blockers, dilators and other devices is crucial to success, and ultimately, patient outcomes. The steady growth and refinement of endoscopes used to see within the airway has developed from simple tubes and reflected light-sources to sophisticated hybrid devices using fibre-optics and integrated video camera systems. A basic understanding of the physical properties that underlie the construction and function of these devices is essential to proper use and care.

These notes aim to give a brief introduction to the important principles, and a broad overview of the main types of endoscopic airway equipment in use. More information, training materials and video tutorials can be found on the open-access airway education web site, www.openairway.org. Specific queries can be directed to ross.hofmeyr@uct.ac.za

Inclusion of images or mention of specific products by name in these notes is for the purposes of discussion and is not an endorsement of the product.

Basic principles of conventional optics

Optics is that branch of physics which describes the characteristics and behaviour of light, instruments that use, detect or manipulate light, and its interactions with other forms of matter. Conventionally, it refers to visible, infrared and ultraviolet light, but as light is a form of electromagnetic wave, it has implications for other forms such as x-rays, radio and microwaves.

Conventional *geometric optics* describe the phenomena that can be accounted for using the classical electromagnetic ray form of light, which presumes travel in straight lines and reflection or refraction when meeting or passing through surfaces. Wave and particle effects such as interference and diffraction are described by the more comprehensive *physical* and *quantum optics,* but are largely beyond the scope of these notes.

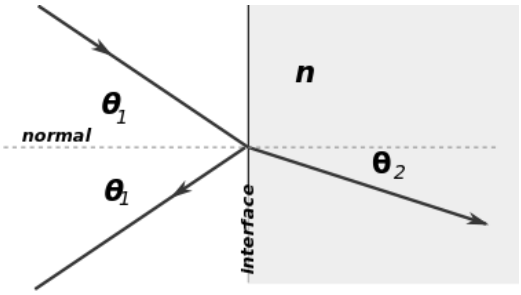


Figure 1. Geometry of reflection and refraction.

The fundamental principles of geometric optics are the laws of **reflection** and **refraction**. Angles of incidence, reflection and refraction are always measured from the **normal**, which is perfectly perpendicular to the interface. When a ray of light meets the interface between two transparent materials, it is split into reflected and refracted rays, so that:

> The reflected ray lies in the plane of incidence, and the angle of reflection equals the angle of incidence (**Law of Reflection)**
>
> The refracted ray lies in the plane of incidence, and the sine of the angle of refraction divided by the sine of the angle of incidence is a constant for any two materials and a given wavelength of light. (**Law of Refraction)**

This can also be expressed as **Snell's Law**, which describes the angles from normal for a light ray traversing from a medium with refractive index of n_1 to a medium with index n_2:

$$\boldsymbol{n_1 \sin \theta_1 = n_2 \sin \theta_2}$$

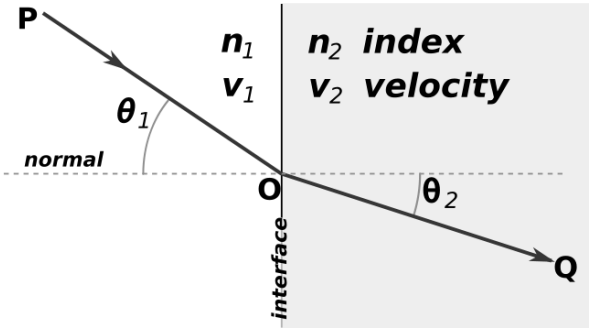


Figure 2. Snell's Law of refraction.

The velocity *v* of light in a transparent medium is determined by the nature of the medium, but is always less than the speed of light through a vacuum. The refractive index *n* for any given medium is therefore calculated as

$$\boldsymbol{n = c/v}$$

where *c* is the speed of light through a vacuum. Where there is a marked difference between the indexes of refraction from one medium to another (such as between glass and air), Snell's law predicts that there is no θ_2 when θ_1 is sufficiently large. In other words, when the angle of incidence is sufficiently far from the normal, no light is refracted, and **total internal reflection** occurs. This is the fundamental principle of the function of fibreoptics (see below).

Lenses are devices which cause light rays to converge or diverge through refraction. Converging lenses focus incoming parallel rays onto a spot on *focal length* from the lens, on the opposite side. Diverging lenses spread the incoming parallel rays in such a way that they appear to have originated at a position one focal length on the same side as the origin.

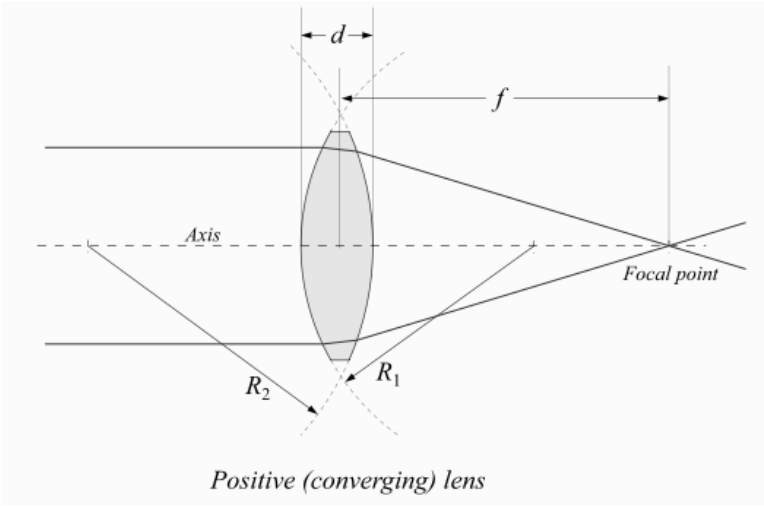


Figure 3. Converging lens. Image: DrBob, creative commons.

Several types of indirect laryngoscope (whereby an image of the vocal cords is produced without a direct visual axis) make used of a combination of lenses, mirrors and/or prisms. Examples include the TruView and Airtraq optical laryngoscopes.

Principles of fibreoptics

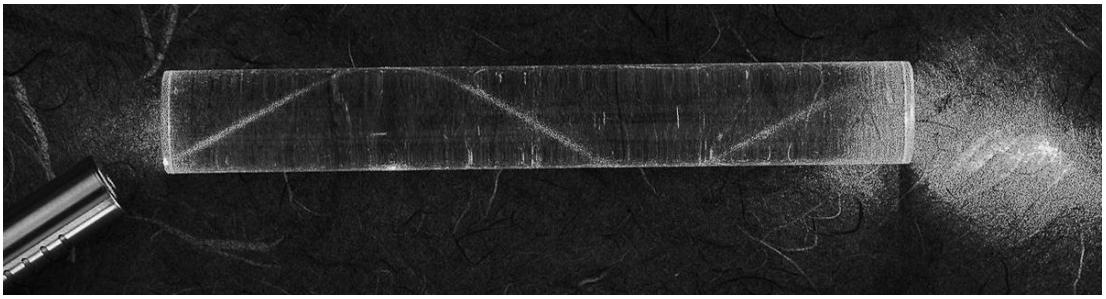


Figure 4. A laser passing down a Perspex rod illustrating the principle of total internal reflection of light in an optical fibre. Source: Wikipedia (User Timwether, creative commons license).

Optical fibres are flexible structures created by sequentially drawing transparent glass or plastic into a very fine diameter, often comparable to or thinner than a single human hair. Using the principle of total internal reflection, they allow transmission of light from one end to the other, potentially over long distances, with minimal loss of intensity. This is referred to as acting as a waveguide. Often, cladding material with a lower index of refraction is used around the optical fibre to further increase efficiency. They have wide-ranging uses in illumination, imaging, sensing and communications.

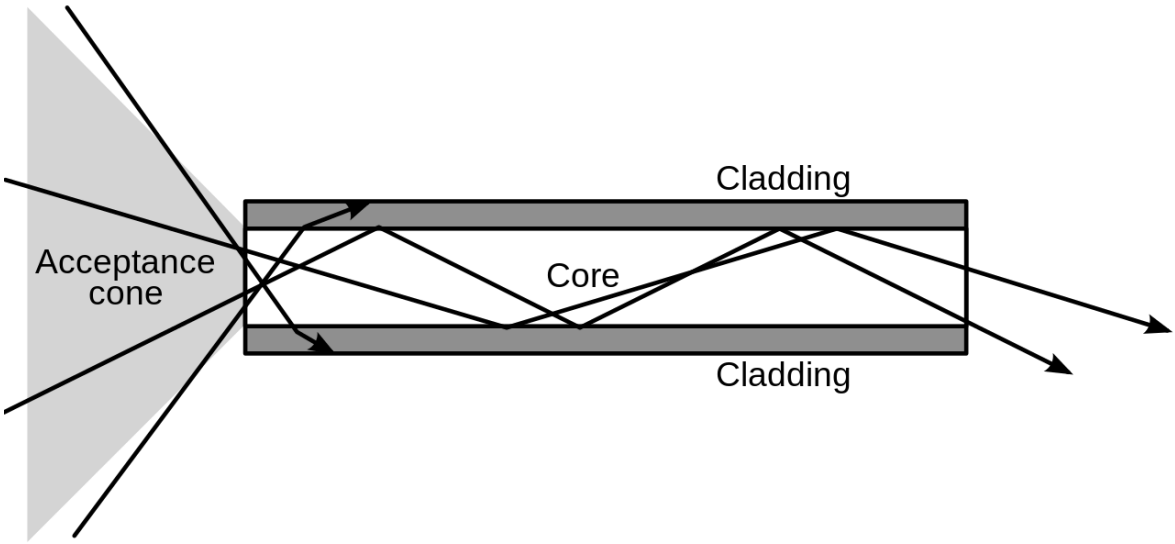


Figure 5. Schematic diagram of an optical fibre. Source: Wikipedia, public domain.

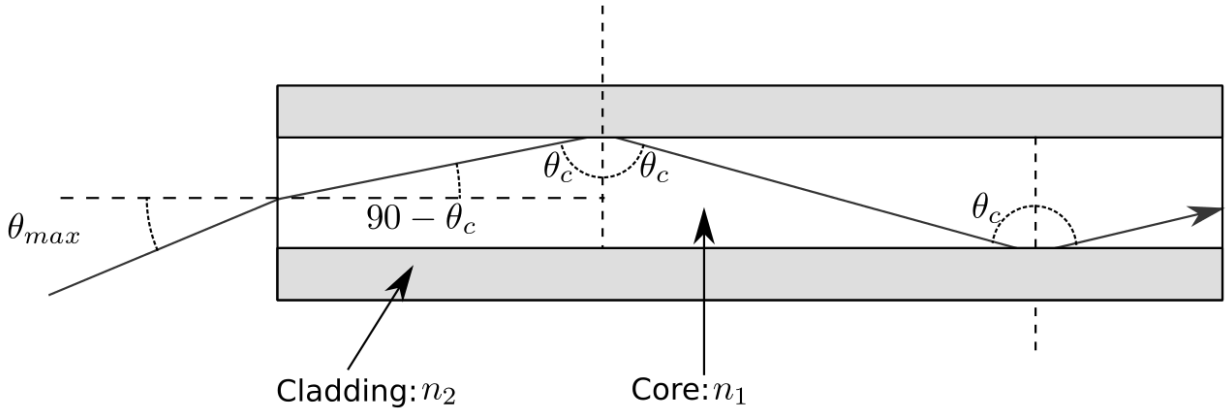


Figure 6. Schematic diagram of an optical fibre, showing the numerical aperture. Light undergoes total internal reflection due to Snell's Law. Source: Wikipedia, User A1 (English Wikipedia, creative commons license).

In medical endoscopes, optical fibre bundles (frequently consisting of tens of thousands of individual fibres) are used both to provide illumination and to produce an image. *Non-coherent* fibre bundles convey light from a light source distant from the patient to the end of the scope, thereby avoiding having a bulb at the scope's tip, which would produce heat and can cause burns. To convey an image, *coherent* bundles (where the fibres maintain their same perfect position parallel to each other from one end of the scope to the other) are used. An objective lens at the tip of the endoscope is used to focus the image on the fibre bundle, and an eyepiece at the user's end allows viewing the image.

Optical fibres allow for the creation of flexible endoscopes, and are used in some types of rigid endoscope for either light conduction, or both light and image conduction. A fibreoptic image is easily identifiable by the 'pixelated' appearance, as each individual 'pixel' is an individual fibre. Due to the very fine nature of optical fibres, however, they are susceptible to breaking through tight bends, hard knocks and crushing. Individual broken fibres can be seen in the image as black 'dead' spots. As the number of broken fibres increases, the scope eventually becomes unusable.

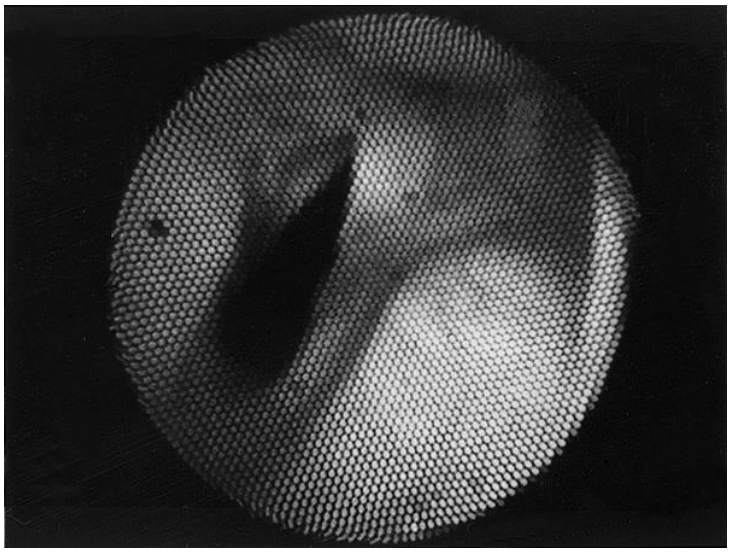


Figure 7. Fibreoptic image, showing typical honeycomb appearance and several 'dead' fibres. Image: Dr Takayuki Kitamura.

Video camera and display integration into endoscopes

The value of performing endoscopy is greatly magnified if the image can be enlarged, displayed on a screen to enable multiple simultaneous viewers, and recorded for documentation, teaching and medicolegal purposes. The initial method to achieve this was to mount a still or video camera onto the eyepiece of an optical scope, and link the camera to a display and/or video recorder. Using modern rod lens telescopes in combination with a high definition (HD) camera, very high quality images can be recorded and displayed. One downside to this system, however, is that the weight of the camera can make holding and manipulating the endoscope more difficult and tiring.

Using an HD camera with a flexible fibreoptic scope is less effective, as the resolution is limited by the density of the fibre bundles themselves. This is often visible as a prominent ‘honeycomb’ pattern on the screen, which can make viewing difficult. To counter the honeycomb effect, slightly defocussing the scope, or the use of advanced imaging filters which interpolate between pixels can be used.

An alternate approach is to completely replace the optical fibres in flexible endoscopes with a digital video system. The advent of very small video camera sensors – and rapid improvement in their quality combined with reduction in cost, largely fuelled by the digital camera and mobile phone industries – has made these ‘chip-in-tip’ endoscopes increasingly common. The light source and non-coherent fibre bundles are also replaced by one or more light-emitting diodes (LEDs) at the end of the scope, which makes flexible video endoscopes lighter, more robust, and either thinner in diameter, or the same diameter with a larger working channel.

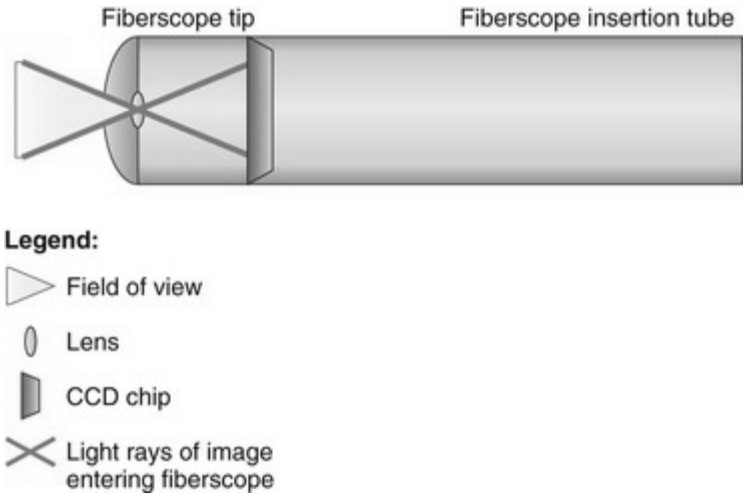


Figure 8. Schematic representation of a flexible video endoscope with 'chip-in-tip' CCD camera.

The original video chips were of the CCD (charge-coupled device) design. A thin silicon wafer is divided into a geometric array of light-sensitive regions (picture elements, or 'pixels') that locally store a charge dependent on the degree of light exposure. After exposure, the charge accumulated by each pixel is transferred across the array in sequence, creating a digital signal which encodes the image. This has been likened to an array of buckets collecting rain:

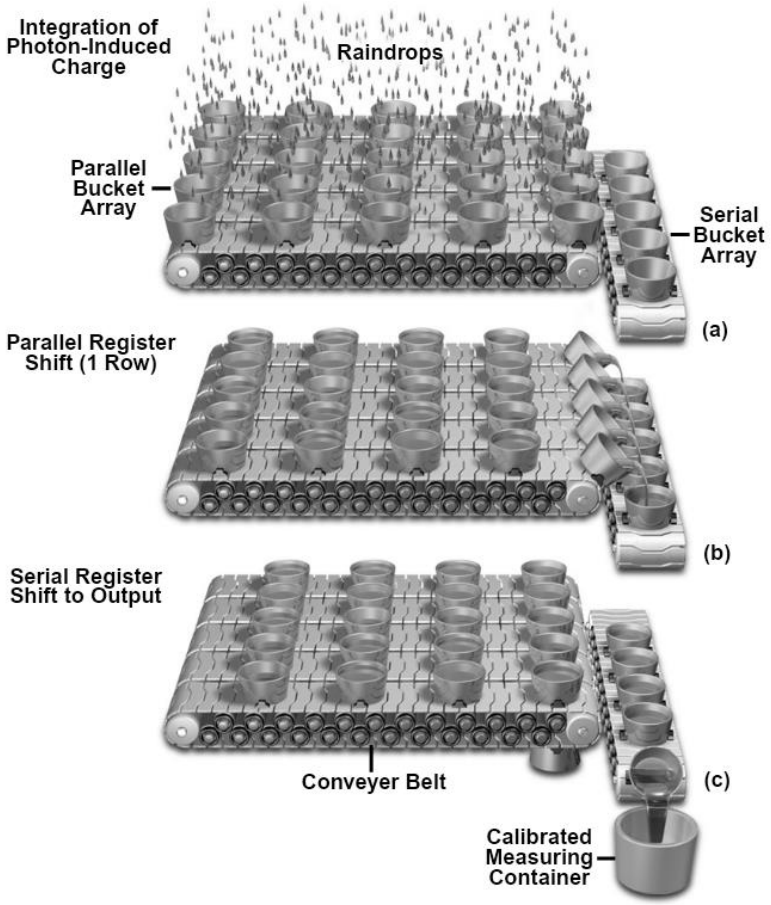


Figure 9. CCD 'Bucket Brigade' analogy: Individual pixels (buckets) collect incoming photons (raindrops) in the form of electrical charge. This is then transferred across the array in orderly steps, and recorded sequentially into a digital signal.

A second type of digital image chip is the CMOS (complementary metal-oxide semiconductor) sensor, which employs a similar array architecture but allows multi-channel recording from the matrix, greatly increasing image capture speed. CCD sensors are more sensitive to light than CMOS, and therefore produce images with less 'noise' at the same light intensity than CMOS. However, in the light-abundant environment of airway endoscopy, this is much less of a disadvantage. CMOS, however, are more power-efficient and can be produced at much lower cost. Steady improvements in CMOS sensor quality has led to their domination of the camera chip market, and most modern video endoscopes use a CMOS chip.

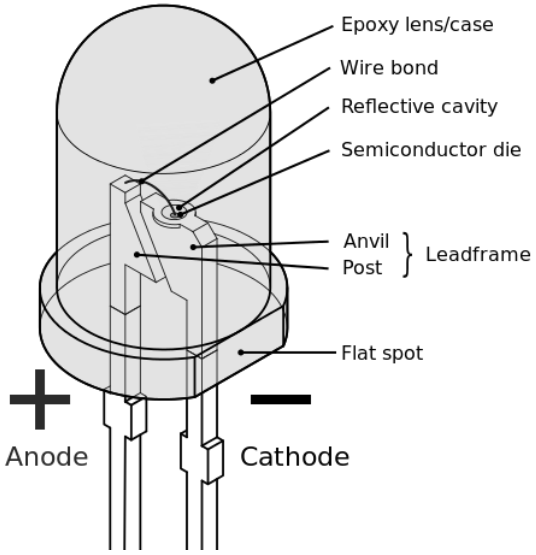


Figure 10. Schematic diagram of a light-emitting diode

Light-emitting diodes (LEDs) are electroluminescent light sources, in which the interface between two semiconductor materials produces photons when subjected to a suitable voltage. They have multiple advantages for the use in endoscopes of all types when compared to incandescent and fibreoptic light sources: small size (less than 1 mm^2, low power consumption, long lifetime (of measured in tens of thousands of hours), physical robustness, and little or no heat production. Although the advent of fibreoptic light guides ushered in the concept of 'cold light sources', only the incorporation of LEDs into endoscopes has truly achieved this goal.

Continually improvement in quality and reduction in both cost and size of both image sensor chips and LEDs has led to an explosion in their use in both flexible endoscopes and video laryngoscopes.

Types of equipment for airway endoscopy

Devices used for airway endoscopy can be classified by the site of desired use (eg. laryngoscopes, bronchoscopes), the method used to transmit an image (eg. direct, standard optics, fibreoptic or video), and specific properties of the design (eg. rigid or flexible). No single system of classification exists for all the available devices. Furthermore, in specific circumstances, equipment from disciplines outside of anaesthesia is often used in complex airway procedures (for instance, rigid telescopes or suspension laryngoscopes). In these notes, various types of direct laryngoscope are not discussed, as this information is broadly available elsewhere.

Figure 11. A structural and functional classification of indirect laryngoscopic devices

Laryngoscopes

Optical laryngoscopes

Optical laryngoscopes use lenses, prisms and/or mirrors to convey the image to an eyepiece, creating an indirect view of the larynx, but allowing guided intubation. They often have the advantage of being either low cost and disposable, or fairly compact and robust. All of the current devices on the market can have some form of camera connected to the eyepiece to allow display on a screen, forming a hybrid optical/video device.

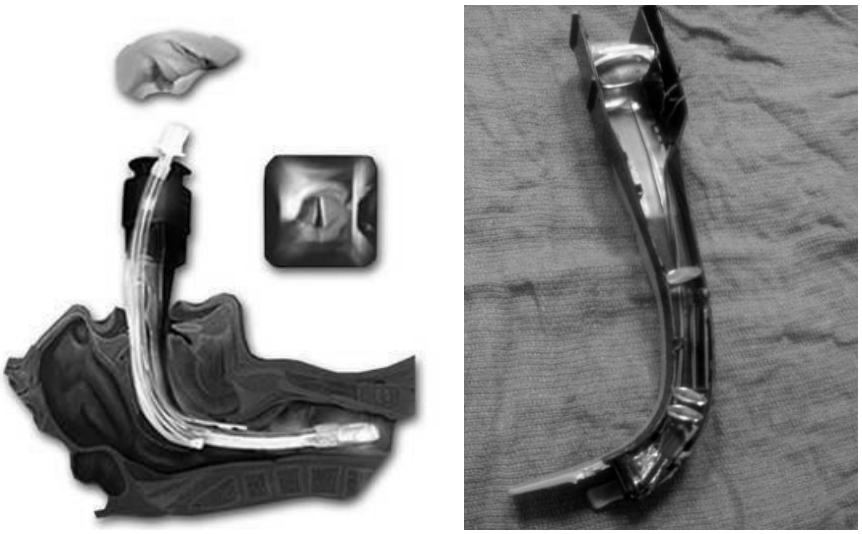


Figure 12(a) and (b). Airtraq optical laryngoscope, showing use and internal construction of lenses, with electrical wires for the LED light source. Note the channelled blade shape. Images: Manufacturer.

Figure 13. Truview optical laryngoscope. Image: Manufacturer

Video laryngoscopes

The number of video laryngoscope (VL) devices on the market has been rapidly increasing. All use some form of CMOS or CCD camera chip and LED light source, and display the image either on a separate screen or display mounted on the handle. Both disposable and reusable blade devices exist. VLs can be classified according to their blade shape into three groups, which dictates their strengths, weaknesses and particular utility in different airway situations:

- Hyperangulated blades, such as the classical Glidescope blade or CMAC D-blade
- Traditional Macintosh or Miller shaped blades
- Channelled blades, such as the Pentax AWS or King Vision VL

Figure 14. A set of video laryngoscopes, showing hyperangulated and conventional Macintosh and Miller blade shapes with a separate video display. Image: Manufacturer

Figure 15. King Vision VL with channelled blades and display mounted on handle.

Intubating endoscopes

Rigid intubating endoscopes

Also known as optical or video stylets, these devices are preloaded with an endotracheal tube and provide a 'through the tube' view during intubation. They are designed to be used alone, or in conjunction with a laryngoscope (known as *dual endoscopy).* Common examples include the Bonfils, Shikani and Levitan optical stylets, and the Clarus and CMAC-VS video stylets. The optical stylets use fibreoptic bundles for illumination and image creation, but are much more robust than flexible fibreoptic scopes due to their rigid construction. Despite this, they are manufactured in sizes down to an external diameter of 2 mm, allowing paediatric tubes of as small as 2.5 mm ID to be used. These devices have the further advantage in very soiled airways of being able to be used as a lightwand if intubation under vision is not possible.

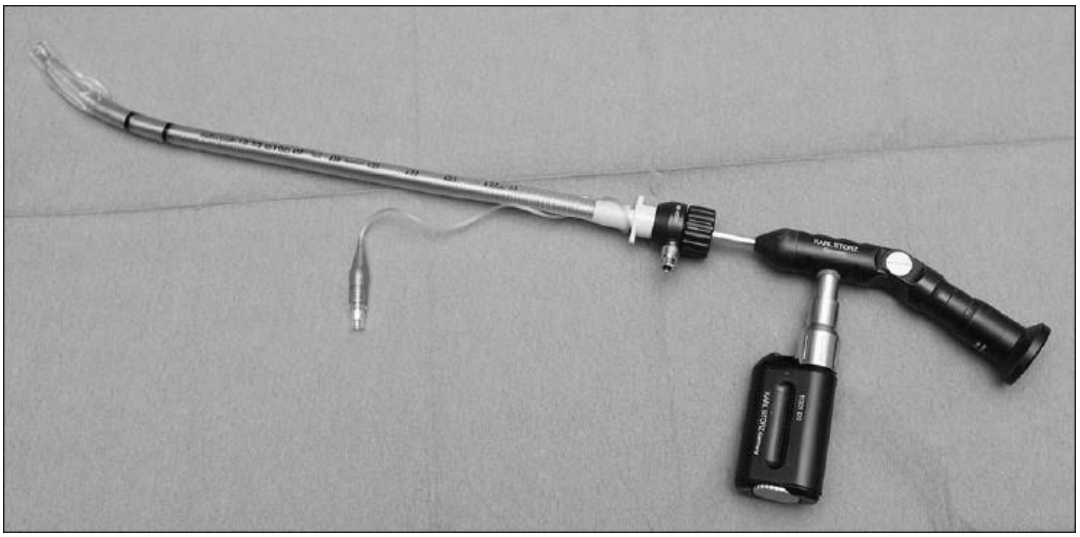


Figure 16. Bonfils rigid intubating endoscope with preloaded endotracheal tube and batter light source.

Flexible intubating endoscopes

Flexible intubating endoscopes are either fibreoptic or 'chip-in-tip' video endoscopes, and are designed for awake intubation of very challenging airways. Functionally, they bear very close resemblance to flexible bronchoscopes (see below), and are often used interchangeably. Modern models have an interface between the control body and insertion tube (or an additional accessory) to securely hold a pre-loaded endotracheal tube during endoscopy for intubation. Typically, the minimum internal diameter of the ETT should be no less than 1 mm greater than diameter of the scope.

Figure 17. Control body and flexible tip of a flexible intubating video endoscope (FIVE)

Bronchoscopes

Rigid bronchoscopes

Essentially a straight metal tube with connections to allow passage of light, rigid telescopes and operating instruments, rigid bronchoscopes can be used with the naked eye (with a fibreoptic light source connected directly to the scope with a prism) or with a telescopic camera. They are useful to gain access to the airway when swelling or external compression causes collapse or obstruction, and offer a large diameter working area for surgical tasks and removal of foreign bodies. Rigid tracheoscopes are a slightly shorter variant without side holes, which allows ventilation through the scope while working.

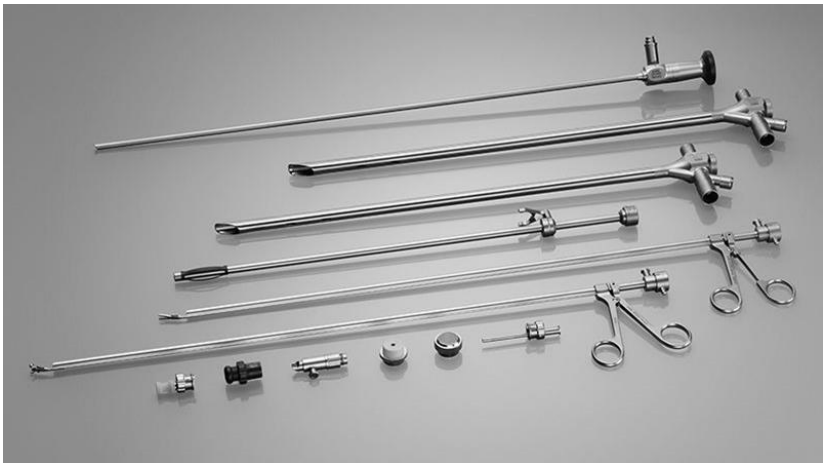


Figure 18. Comprehensive system for rigid bronchoscopy, including rigid telescope, several bronchoscopes, graspers, and attachments for jet and conventional ventilation as well as light delivery.

Flexible bronchoscopes

Practically speaking, there are trivial differences between flexible intubating endoscopes and flexible bronchoscopes. The former are usually slightly (~5 cm) shorter and slimmer (5.0 – 5.5 mm external diameter), and the latter place larger emphasis on having a larger working channel for graspers and biopsy forceps at the expense of a slight increase in diameter (5.8 – 6.3 mm). 'Paediatric' versions of both scopes exist, with external diameters of usually 3.0 – 4.0 mm. However, taking into account the implications of scope diameter for endotracheal tube size selection, they function perfectly well for intubation.

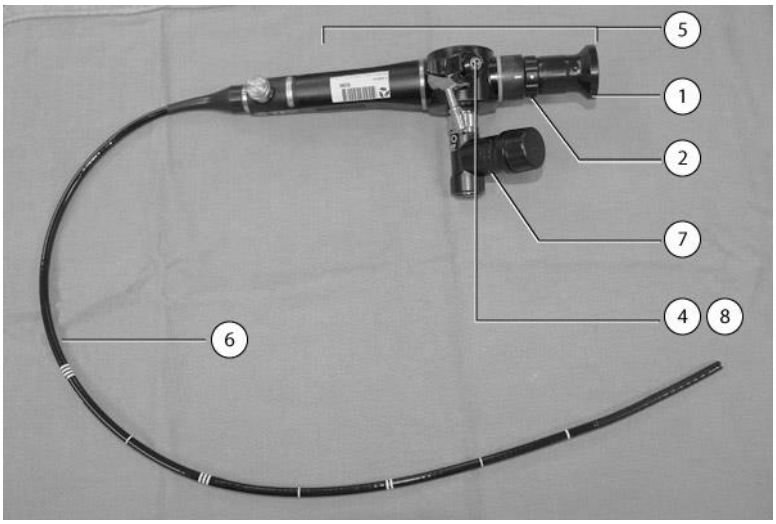


Figure 19. Flexible fibreoptic bronchoscope. Parts: 1) Eyepiece 2) Focus ring 3) Control lever 4) Working channel port 5) Control body 6) Insertion cord 7) Light source 8) Suction valve/port. Image: Toronto General Hospital Department of Anaesthesia online learning

Figure 20. A portable flexible video bronchoscope with integrated display. Image: Manufacturer

Operative telescopes

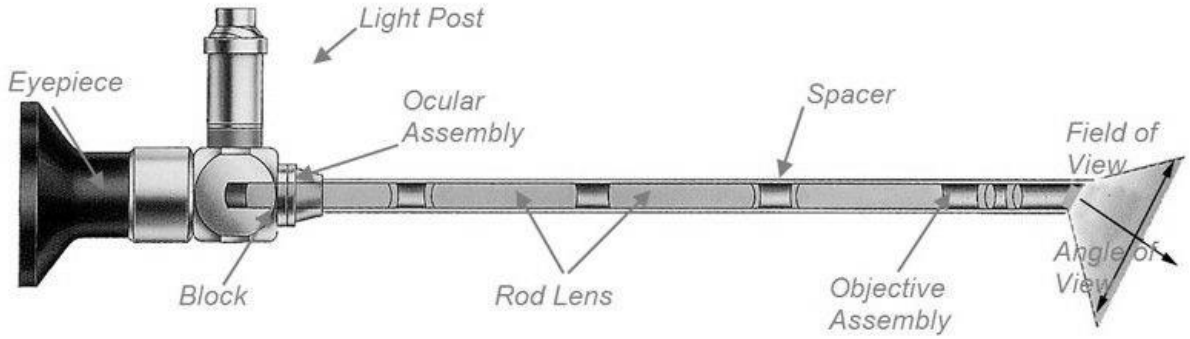


Figure 21. Typical construction of an operative endoscopic telescope, utilising a rod lens design. Image courtesy Fuzhou Alpha Optics.

Originally, endoscopic telescopes had a traditional lens design, which featured small lenses separated by large air gaps. Physicist Harold Hopkins realised, however, that an alternate design in which the lenses where long rods separated by small airspaces was more efficient, and did not require structures to hold the lenses in place, increasing their size and optical efficiency. This increased the image quality, brightness, and field of vision, while at the same time making the endoscope more robust. After patenting his design in 1959, he was approached by the German optical instrument maker, Herr. Karl Storz, to incorporate the design into his endoscopes. This was a turning point for both men; Hopkins became famous for the design (amongst others), and Storz grew his company into the foremost manufacturer of rigid endoscopes in the world. Most modern operative telescopes now use this design, which can be used to create scopes with diameters as small as 1.2 mm while still producing images far better than those achieved with today's chip-in-tip sensors.

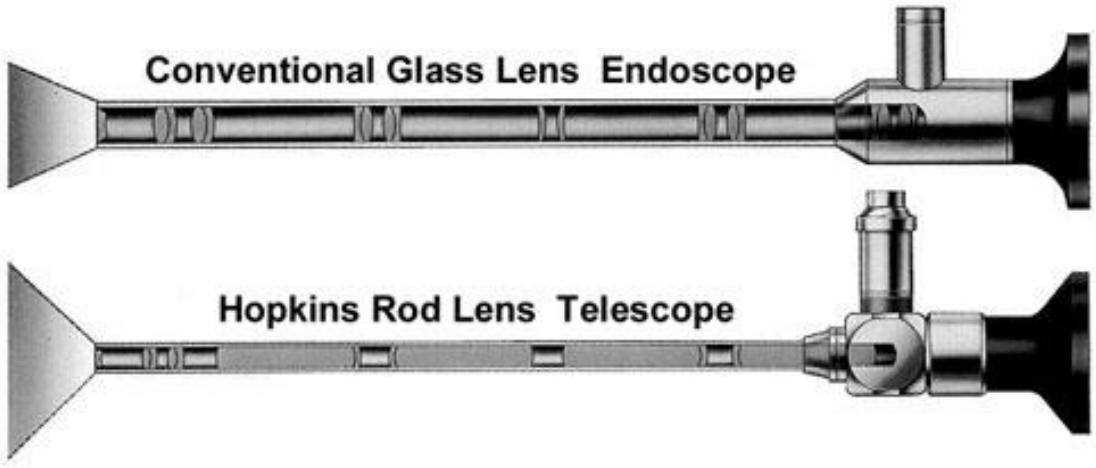


Figure 22. Conventional operative telescope (endoscope) in comparison with the Hopkins 'rod lens' design. Image courtesy of Karl Storz GmBH.

The Future

Clearly, endoscopic airway devices are continuing to evolve, and with them, our surgical and anaesthetic techniques. Patients previously deemed very difficult airways for whom only an awake fibreoptic intubation was advocated are now routinely intubated with asleep video laryngoscopy, and procedures such as tracheal dilatation which were only performed with rigid bronchoscopes and bougies are now being achieved as flexible endoscopic day cases through a supraglottic airway. Just as many procedures in cardiac surgery are not achievable without an ‘echo anaesthetist’ who can provide intraoperative echocardiographic views, so too are procedures in ENT and thoracic surgery beginning to require an ‘interventional anaesthetist’ who can both control and image the airway. However, quantifying the contribution of these novel and advanced techniques to a wide base of patient outcomes remains a challenge. It is incumbent upon the individual practitioner to become highly skilled with the tools at their disposal, and to recognise the relative strengths and weaknesses of each device in each situation.

Key Concepts in Understanding Vaporizers

Dr Rowan Duys

UCT Dept of Anaesthesia & Perioperative Medicine
UCT-Dräger Fellow in Education and Simulation

Introduction

In recent times, the mechanisms employed in the design and manufacture of anaesthesia vaporizers have evolved only through the incorporation of electronic processing units to increase the accuracy and safety of the measured content of volatile delivered.

However, this fact has not dampened the examiners' appetites for interrogating candidates' knowledge of the physical properties that govern their design.

This review will not attempt to re-invent the wheel. The 2011 CEACCP by Boumphrey (link: http://tinyurl.com/mfzrjtd), and the ARC notes by Khalpey 2015, and Reid 2013, complement the chapter on vaporizers in Basic Physics and Measurement in Anaesthesia by Davis and Kenny. These cover what would appear to be sufficient material for the Part 1 candidate (note: I wrote my Part 1's in the UK, and I am not, nor never will be, a Part 1 examiner…so my insight is limited…but surely SURELY this is enough?)

However, one new concept has emerged through a fascinating editorial by James, Hofmeyr and Grocott (link: http://tinyurl.com/kgubvcn) that discusses an apparent error in how vaporizers are calibrated and vapour content measured and reported by our anaesthesia monitors. Apparently we've been doing it wrong…forever!

I will attempt to describe the relevant physics at work in our vaporizers, and briefly describe the conflict between expressing volatile content as a concentration when we actually mean a partial pressure. For a more detailed description of the workings of a vaporizer, please view the references mentioned above.

Important Definitions

Gases and Vapours

Critical Temperature is the unique temperature of a substance, beyond which it cannot be transformed from a gaseous state to a liquid state, no matter how much pressure is applied. We call substances above their critical temperatures gases.
When a substance exists in a gaseous form at or below its critical temperature, it is known as a **vapour.**

Saturated Vapour Pressure

In any liquid, some molecules will acquire the energy required to break the tight bonds of the liquid phase and escape as a vapour. *(They will be free to explore the world on mad-cap adventures like going out for dinner; they probably do not have children.)** This process is known as **evaporation** and requires heat energy known as the **latent heat of vaporization**. It only occurs at the surface of the liquid.

Equally some molecules in the vapour form will return to the less high energy state and become liquid again. *(These molecules, no-longer capable of roaming free to enjoy themselves, probably have several children)**

If a liquid is placed in a sealed container a dynamic equilibrium will be reached where molecules become vapour and return into liquid phase at an equal rate. This vapour state is therefore said to be saturated and the pressure exerted by this vapour is considered the **saturated vapour pressure** (SVP). The SVP changes as the temperature of the liquid changes, but is independent of the ambient pressure.

At sea level, the SVP of sevoflurane in a closed container (like a vaporizer) at 20˚C is 21,3 kPa. If you reduce the ambient pressure by travelling to altitude, but the temperature remains 20˚C, the SVP will

stay the same. However, if you increase the temperature of the liquid phase of Sevoflurane, the SVP will increase.

Boiling

The SVP for a particular liquid increases with temperature in a non-linear fashion. As the temperature rises, the SVP will reach a point where it equals the atmospheric pressure, **the boiling point**, which will allow molecules to undergo a phase change to vapour not just at the surface, but deep within the liquid. This forms bubbles of saturated vapour that rise to the top in a process known as boiling.

Vaporizers

A vaporizer is a device for reliably delivering clinically useful, accurate and adjustable amounts of volatile anaesthetic vapour to a carrier gas.

If the carrier gas leaving the vaporizing chamber of a vaporizer is fully saturated, the concentration of the vapour can be calculated using the following equation:

$$gas\ concentration = \frac{vapour\ pressure}{ambient\ pressure}$$

So in the case of Sevoflurane at 20˚C:

$$gas\ concentration = \frac{21{,}3kPa}{101{,}3kPa} = 21\%$$

21% is not a clinically useful concentration and thus vaporizers are designed to dilute the amount of vapour in a carrier gas by addition of fresh gas. And two methods exist for doing this:

- **Variable flow vaporizers** adjust the proportion (or splitting ratio) of fresh gas that flows through or bypasses the vaporizing chamber.

- **Measured flow vaporizers** add vapour directly to the fresh gas stream and this group of vaporizers include the Ohmeda Tec 6 designed for use with Desflurane, and direct injection of volatile anaesthetic (DIVA) vaporizers

I don't believe I can add meaningfully to the description of the workings of the various vaporizers present in the reference list and the science has not changed meaningfully since the textbooks were written.

Potential problems with vaporization

Adequate vaporizer function depends on saturation of carrier gas leaving the vaporizing chamber. Problems with this function are now largely historical due to modifications in vaporizer design but the applied physics inherent in their description keeps them high up in the examiner's 'go-to' list. Reid's ARC from 2013 explores the issues in more detail but they are summarized:

Poor Vaporization

Vaporization rates are improved with increased surface area and various mechanisms are employed to maximize this including baffles, wicks and bubbles. Cowls are also employed to direct gas flow over the surface of the volatile.

Temperature Fluctuation

SVP is temperature dependent and evaporation leads to cooling, so the SVP will reduce as evaporation occurs and temperature drops. The two mechanisms employed to limit this are attempts at stabilizing the temperature of the liquid with the use of materials of high heat capacity and thermal conductivity, and temperature regulated adjustments to the splitting ratio achieved with devices that alter their shape and thus inlet aperture size with temperature changes.

Flow dependence

High fresh gas flows reduce the ability to saturate carrier gas in the vaporizing chamber and can cause preferential flow through the lower resistance bypass channel. Devices to improve vaporization and ensuring equal resistance of the bypass and vaporizing channels have reduced this.

Back Pressure

Pressure in the conduit from vaporizer to circuit, created by positive pressure ventilation can alter the amount of vapour present in the carrier gas by a 'pumping effect' where saturated gas is forced backwards out of the vaporizing chamber and into the bypass channel, or by a 'pressure effect' where the ambient pressure in the vaporizing chamber is increased, thus reducing the effective concentration of vapour in the outlet of this chamber ($gas\ concentration = (SVP)/(ambient\ pressure)$).

High internal resistance of the vaporizers (as in Plenums), equal volume bypass and vaporizing channels, long inlet/outlet conduits and one-way valves have reduced these phenomena.

Fresh gas composition

A clinically insignificant, though examinationally important *(examinationally is a word, I promise)**, effect of using high N_2O concentrations in the carrier gas may be large volumes of N_2O going into solution in the volatile, thus altering the volume of carrier gas leaving the vaporizing chamber, and the resultant concentration of vapour. Differing viscosities may also cause a theoretical alteration in splitting ratios.

Altitude

Changes in ambient pressure due to altitude do not change the SVP of the anaesthesia agents, and thus the partial pressure of the delivered vapour will not change. However, the concentration will. Practically, a vaporizer set to 1% will deliver the same partial pressure of agent whether at sea level or altitude, despite the delivered concentration changing. Put another way, despite vaporizers being marked with concentrations (1%, 2% etc.), the concentration they deliver changes with ambient pressure, but the partial pressure remains the same.

This is fortunate, because what matters clinically, or what actually affects depth of anaesthesia is the delivered partial pressure of the agent, not the concentration. So, when a vaporizer is set to 1% in Cape Town, the concentration delivered will be 1%, and the partial pressure delivered will be 1kPa (1% of 100kPa, the ambient pressure). However, when the vaporizer is set to 1% in Johannesburg at an altitude of 1750m and an approximate ambient pressure of 80kPa, the concentration delivered will be described by the equation: $C_1P_1 = C_2P_2$. Therefore:

$$C_{JHB} = \frac{C_{CPT}.P_{CPT}}{P_{JHB}} = \frac{1\%.100kPa}{80kPa} = 1{,}25\%$$

But the partial pressure, 1,25% of 80kPa will remain 1kPa

Does anybody else think that's weird? When you adjust the concentration dial on the vaporizer in Johannesburg, the actual concentration of the delivered vapour as measured by the gas analyser, will be different, but the partial pressure that results will be what we want it to be? Well, luckily for us, James et al, also think it's weird, and we're getting there. But calm down, chill your boots, we need to discuss Desflurane vaporizers first.

Desflurane

The SVP of Desflurane at 20˚C is 88,5kPa. The fresh gas flows required to dilute this to a clinically useful partial pressure using a bypass channel in a variable flow vaporizer would be too high to remain economical. Also, the boiling point of Desflurane is about 23˚C resulting in intermittent boiling at room temperature and alterations in delivered content.

Desflurane vaporizers are thus different to normal variable flow vaporizers. They are heated to 39˚C to set the SVP at around 2atm and various electronically controlled mechanisms are used to measure the flow to the patient and add the correct amount of 100% Desflurane to the fresh gas. Note, 100%

Desflurane is added to a measured fresh gas flow (hence **measured flow vaporizer** or gas/vapour blender) vs variable amounts of fresh gas being diverted through a bypass channel and vaporizing chamber (**variable flow vaporizer**)

See the reference texts for more details on the Desflurane vaporizer and the Aladin cassette, both of which are examinable.

Safety features

It is worthwhile understanding the various mechanisms employed to ensure safe delivery of volatile agents. They are well described in Reid's refresher notes.

MAPP vs MAC and why we've been doing it wrong

Losing concentration: time for a new MAPP is a BJA editorial from 2015 with 2 South Africans and a big-cheese in altitude medicine (think the landmark Everest NEJM PaO_2 study) as co-authors. The last time I can remember a big BJA editorial with South African authors it ended up in my Part 2 orals #justsaying.

What the article Boyles** down to is that all variable bypass vaporizers are designed to deliver **varying dilutions of SVP**. They thus **deliver a partial pressure** not a concentration.

Further, all gas analysers actually **measure partial pressure.** The fact that anaesthesia machines and monitors display end-tidal concentrations is largely historical, dating from a time when only volumetric measurement of gas was possible. And thus partial pressure measurements are converted to concentrations or percentages before reporting.

And lastly, that **partial pressure is the determinant of depth of anaesthesia**, not concentration.

What this translates to with the current nomenclature and calibration is a risk of variable partial pressures of agents being delivered at different altitudes with resultant problems of awareness and compromise to patient safety.

James et al propose a change in all anaesthesia machines to report the measured **partial pressure**, and not the derived value for concentration as a means to reduce risk and standardize practice.

As stated previously, I am not a Part 1 examiner, and so actually have no idea if this work is examinable or not, but it is one of very few articles published about vaporizers in the last 3-5 years that brings something novel to the discussion. The concepts of altitude and changes in agent delivery are well explained. Plus, it's not often a Part 1 article is also a fun read 😐….do it!

*if it's in italics, I made it up, and it has nothing to do with vaporizers, I promise, and apologise
**I'm here all week

Please do contact me if you'd like to discuss any points raised here or are struggling with access to the references. rowanduys@gmail.com

References

1. Reid; Understanding Vaporizers; Part 1 ARC notes 2013
2. Khalpey; Vaporizers, Part 1 ARC notes 2015
3. Boumphrey, Marshall; Understanding vaporizers. Contin Educ Anaesth Crit Care Pain 2011; 11 (6): 199-203
4. James, Hofmeyr, Grocott; Losing concentration: time for a new MAPP?. Br J Anaesth 2015; 115 (6): 824-826
5. Davis, Kenny; Basic Physics and Measurement in Anaesthesia; 5th Edition, 2005

Visco-elastic Coagulation Testing

Dr Malcolm Miller

Division of Critical Care
UCT Dept of Anaesthesia & Perioperative Medicine

Introduction

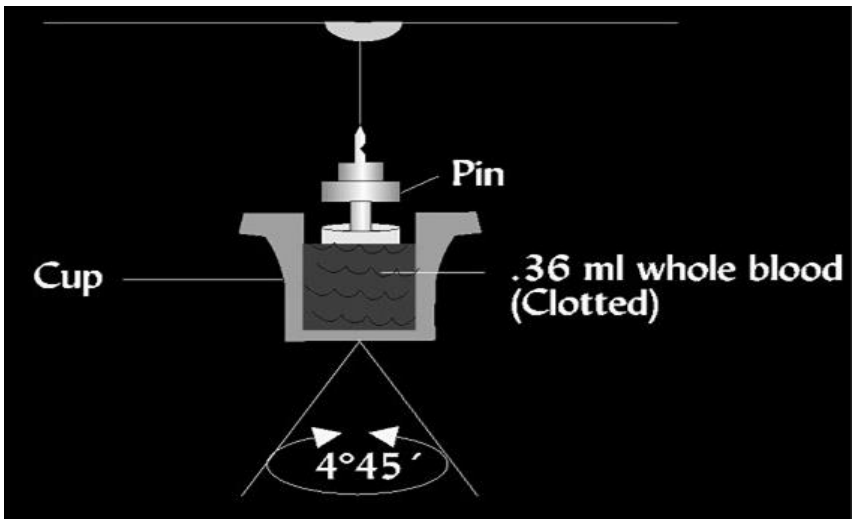
The value of conventional tests which assess coagulation, such as INR and PTT, has been questioned, particularly in the peri-operative setting. These tests are done on plasma and citrated samples, and have to be shuttled off to the laboratory for testing. The resurgence in point of care testing has brought renewed enthusiasm for thromboelastometry, for the assessment of coagulation defects on whole blood samples.

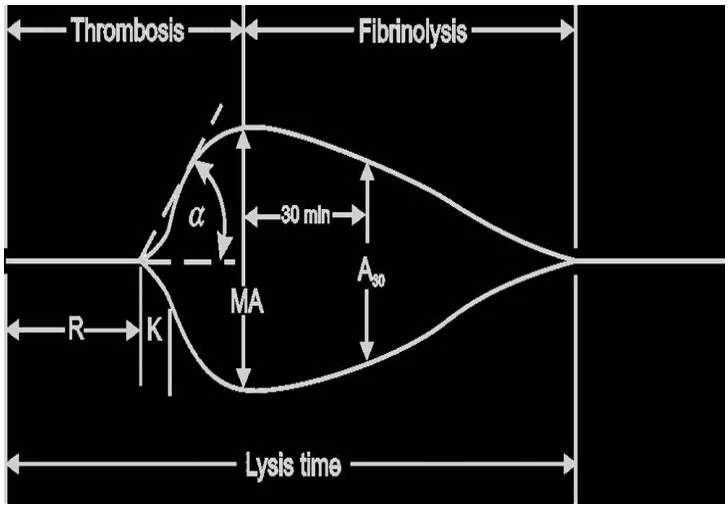
Visco-elastic coagulation devices

Thromboelastography was first described by Hartert in 1948 as a method to assess global haemostatic function from a single blood sample.
These devices assess the visco-elastic properties of blood samples under low shear conditions.
It represents a dynamic global assessment of coagulation from the point of clot initiation to clot stabilization and furthermore evaluates the breakdown of the clot.
The visco-elastic properties of clot formation represent the interaction of fibrin and platelets and is directly related to the strength of the clot formed. The clot then retracts and lyses which is also represented in terms of visco-elastic forces.
Important limitations of these devices are that they assess coagulation under static (no flow) conditions and in a cuvette which negates the effect of the endothelium.

Thromboelastography

TEG® is a registered trademark of Hemoscope Corporation and has been used exclusively to describe coagulation testing using their device.
The TEG measures the clot's visco-elastic properties using a cylindrical cup containing 0.36ml of whole blood that oscillates through an angle of 4^0 45'. Each rotation cycle lasts 10s. A delicate pin in immersed in the cup and the torque of the rotating cup is transmitted to the pin only once fibrin-platelet bonding takes place which links the cup and pin together. The strength of these fibrin-platelet bonds determines the magnitude of the pin motion. Thus the output is directly related to the strength of the clot formed. As the clot lyses, these bonds are broken and the motion transmitted through the pin is diminished. The rotation movement of the pin is converted by a mechanical-electrical transducer to an electrical signal, finally being displayed as a typical TEG tracing.



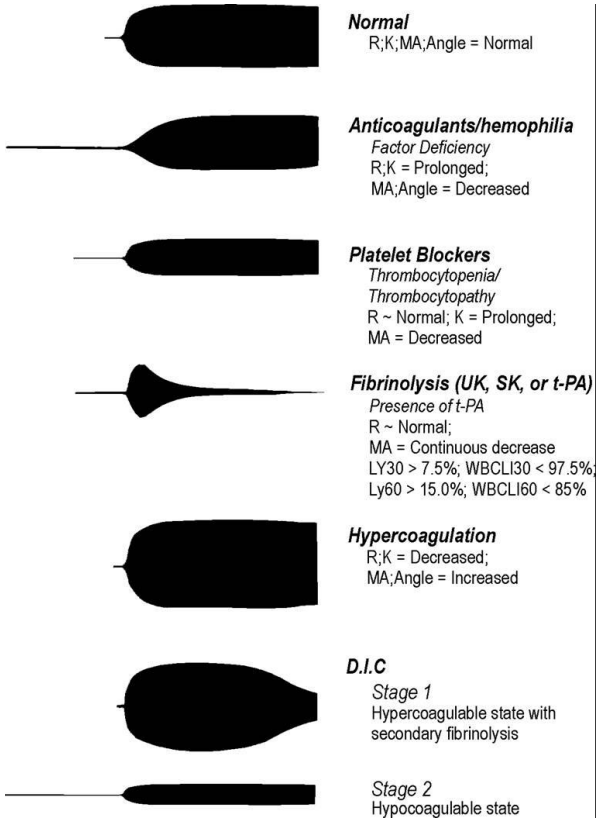


Nomenclature:

R-time: reaction time or time to clot initiation where tracing deviates by 2mm from the baseline
k-time: time from deviation of the tracing from 2mm to 20mm
α-angle: slope between *R* and *k*
MA: maximum amplitude
$MA_{30\text{-}60}$: amplitude at set time
G: clot elasticity (software derived value)
CI: coagulation index, an indication of clot integrity
CL: clot lysis usually at 30min and 60 min after the MA, expressed as a percentage
EPL: estimated percent lysis calculated 30min after the MA

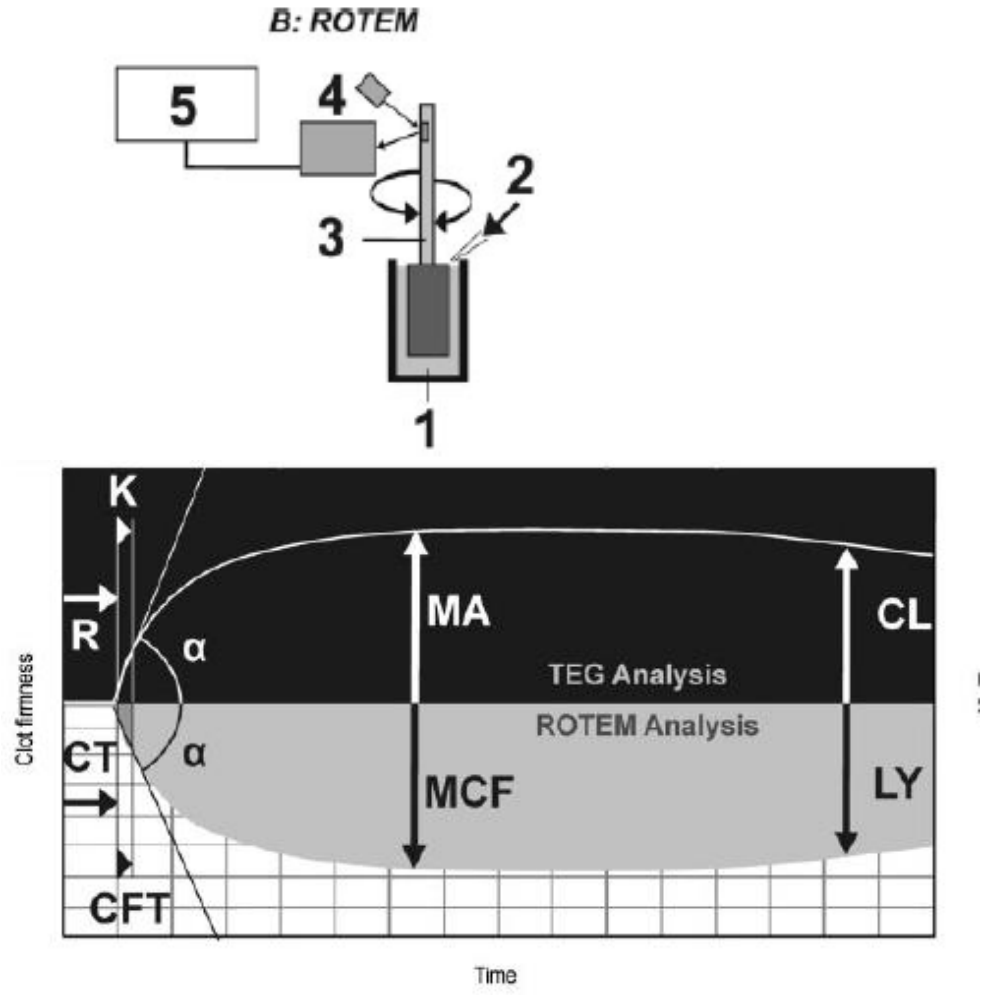
Importantly, TEG provides valuable information regarding both hypocoaguable states as well as prothrombotic states. Remember, clot formation is accelerated by the addition of kaolin to the whole blood sample and if heparin contamination is suspected, this can be neutralized by the use of a cup coated with heparinase.

Examples of TEG tracings:



ROTEM

The device has been commercially marketed as using rotational thromboelastometry, as opposed to thromboelastography used by TEG, but the technology is virtually identical. The ROTEM differs slightly from TEG in that the signal from the pin is transmitted via an optical sensor, not a torsion wire, and the movement is initiated from the pin, not the cup.



ROTEM nomenclature:

CT: clotting time
CFT: clot formation time analogous to r-time plus k-time for TEG
MCF: maximum clot firmness
LY: lysis usually measured at 30min and 60 min

Assays available using ROTEM:

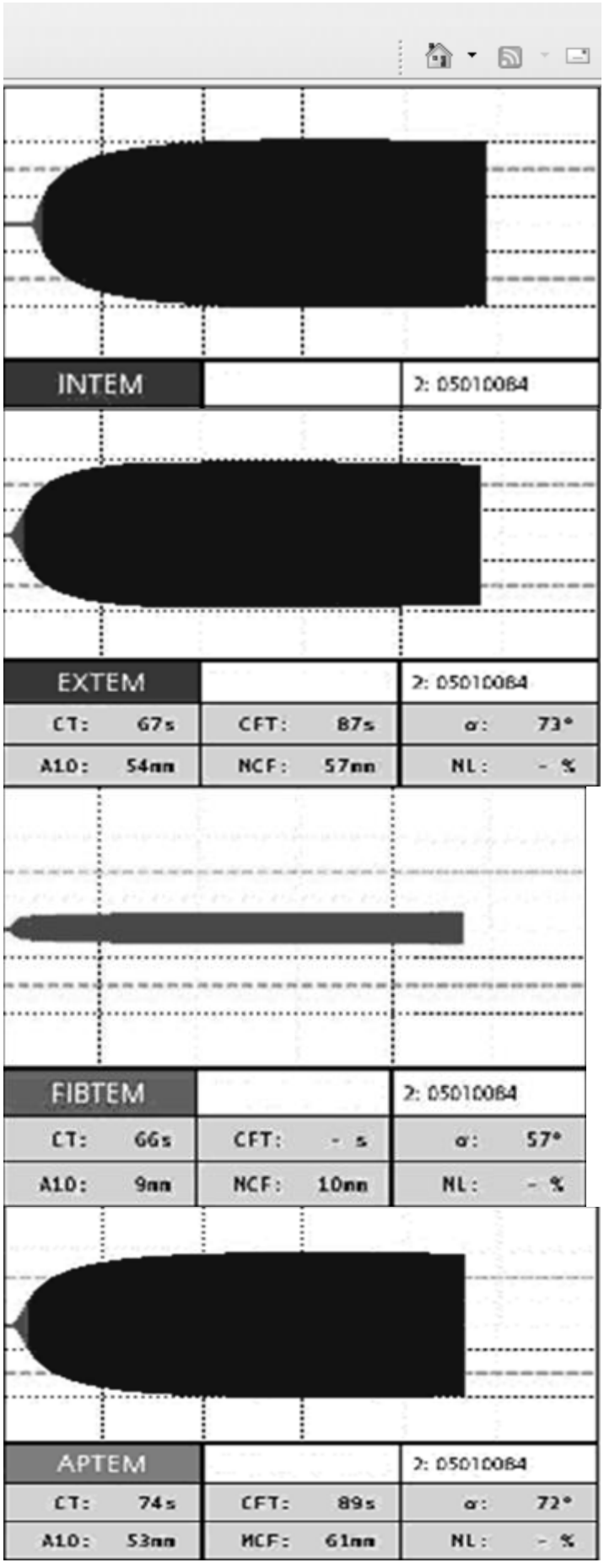
EXTEM: clotting activated by addition of 20µl of tissue factor
Analogous to prothrombin time (PT)

INTEM: clotting activated by phospholipid
Analogous to aPTT

FIBTEM: evaluation of clot integrity by adding cytochalasin D, a platelet inhibitor, evaluating functional fibrin polymerization without platelet activity.

HEPTEM: heparinase added to inactivate the effect of heparin

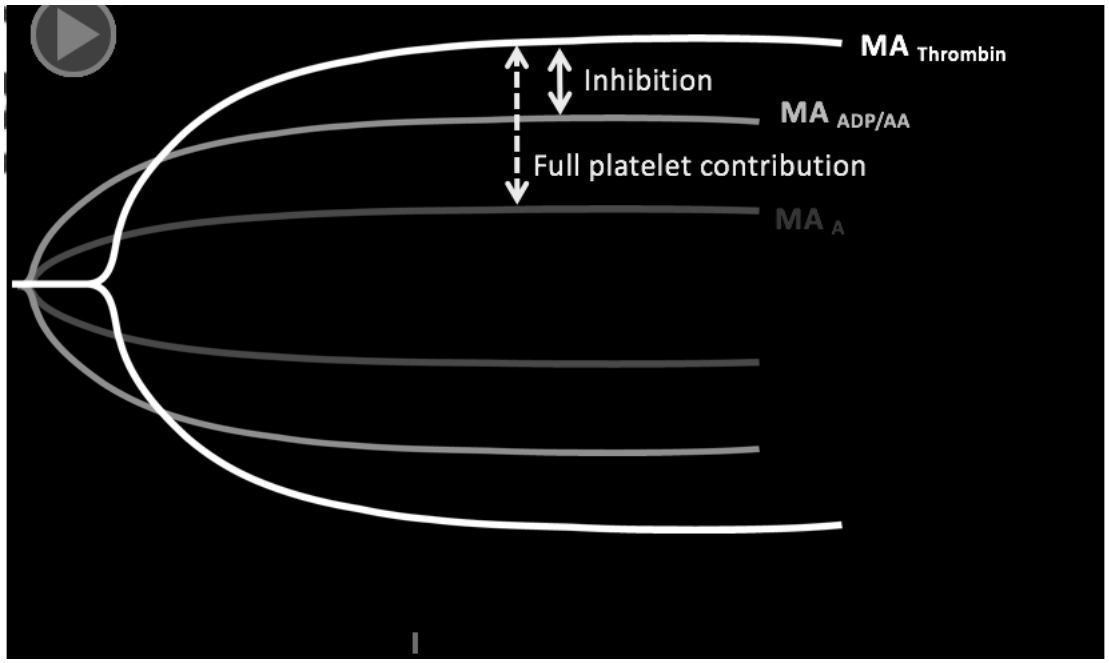
APTEM: aprotinin added to evaluate the presence of fibrinolysis


INTEM
?: 05010084
EXTEM
?: 05010084
CT: 67s
CFT: 87s
α: 73°
A10: 54mm
MCF: 57mm
NL: - %
FIBTEM
?: 05010084
CT: 66s
CFT: - s
α: 57°
A10: 9mm
MCF: 10mm
NL: - %
APTEM
?: 05010084
CT: 74s
CFT: 89s
α: 72°
A10: 53mm
MCF: 61mm
NL: - %

These assays provide a unique opportunity to evaluate the contribution of various factors in the causes of coagulation disorders and offers a distinct advantage over the TEG. However, the test, if all five assays are used comes at a higher cost when compared to the TEG. The device is fully automated and comes with an electronic pipette making operator error less likely.

Platelet mapping

Platelet function is assessed by a variety of methods indicative of the important role that platelets play in the formation of clot. These tests include bleeding time, aggregometry, clot retraction and platelet release markers such as ADP, platelet factor 4 etc.
TEG and ROTEM have recently added platelet mapping to their armamentarium to further evaluate the effects of anti-platelet drugs on clot formation. These involve a series of consecutive tracings that have some platelet antagonist added to the reagents. Depending on what type of platelet blocker the patient is receiving, the blood is supplemented with either ADP or arachidonic acid and the subsequent tracing reconstitutes the coagulation profile to 'normal' and the deviation from the original baseline tracing will indicate the degree of platelet inhibition present in that particular patient.
This information can be useful to elucidate the degree of platelet dysfunction present in patients presenting for surgery or who are actively bleeding. Alternatively, any form of resistance to anti-platelet drugs can be assessed.



Conclusion

The newer point of care coagulation testing devices have emerged as worthy contenders to more conventional time-honoured tests of coagulation. They offer quicker results and might assist clinicians in targeting therapy more accurately when treating complex coagulation disorders. Areas where theses devices have been found to be particularly useful are in the trauma settting, cardiac surgical patients undergoing cardio-pulmonary bypass and in the peri-operative setting where major blood loss has occurred. It has been established as a discriminating test when considering whether patients have a surgical or medical cause for bleeding. Finally, these devices have shown cost benefit when managing patients with bleeding in the above settings by directing therapy and avoiding the empiric use of expensive blood products liberally.

References:

1. N. Salooja and D.J. Perry, *Thromboelastography: Blood coagulation and Fibrinolysis,2001,12:327-337*
2. M. Ganter and C. Hofer, *Coagulation monitoring: Current techniques and Clinical use of visco-elastic Point of Care coagulation devices; Anesth Analg, vol 106, No.5, May 2008.*
3. TEG Hemostasis Analyzer User Manual, Haemoscope Corporation, Niles IL USA.