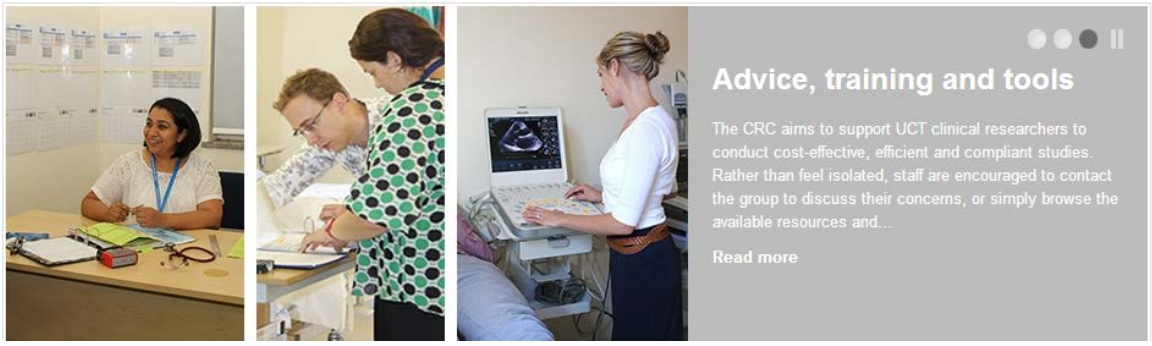
UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

UCT CLINICAL RESEARCH CENTRE
every step of the study

University of Cape Town

Clinical Research Centre

Essentials for the Clinical Research Supervisor Workshop



Faculty of Health Sciences, Clinical Research Centre
"Essentials for the Clinical Research Supervisor"

This workshop is designed to provide supervisors of MMed students' essential guidelines, key milestones and tools for a successful supervisory process. This is a CPD accredited course.

DATE:		
TIME	**TOPIC**	**SPEAKER**
13h30 - 13h45	**Introductions**	Dr Delva Shamley
13h45 - 14h15	**Transitioning from researcher/lecturer to supervisor** Skills, goals, benefits and support team	Dr Celeste de Jager
14h15 - 14h30	**Memorandum of Understanding** Expectations on both sides Postgrad office requirements	Dr Celeste de Jager
14h30 - 14h45	**Intro to FHS Library services** Links will be emailed to you	Dilshaad Brey
14h45 - 15h15	**The Process: Project Designs** "Turn it In"	Dr Celeste de Jager Dr Delva Shamley Sam Lee Pan
15h15 - 15h30	**WORKING TEA BREAK**	
15h30 - 16h00	**What we told your Registrars about Statistics**	Ms Annemie Stewart
16h00 - 16h30	**The Meeting:** Utilising the time/feedback/debate Keeping the momentum Timetable and milestones	Dr Delva Shamley
16h30 - 17h00	**Writing Up**	Dr Delva Shamley
17h00 - 17h30	**Discussion/feedback forms**	All

MMed Supervision

Memorandum of Understanding & Ethics Submission

SUPERVISORS' WORKSHOP

CELESTE DE JAGER
SCHOOL OF PUBLIC HEALTH

Introductions

- Aim of the workshop – clinical aspects

The team:-

- **Clinical Research Centre**- Dr Delva Shamley
- **Biostatistical support** – William Msemburi
- **Epidemiology** (MMed/MPhil) and **Supervisor support** – Dr Celeste de Jager

Transitioning - goals

Clinical Researcher

- Has the qualifications
- Has clinical experience
- Has been through the research process at some level, eg. MMed
- Has ideas to explore
- Funding for projects
- Networking experience

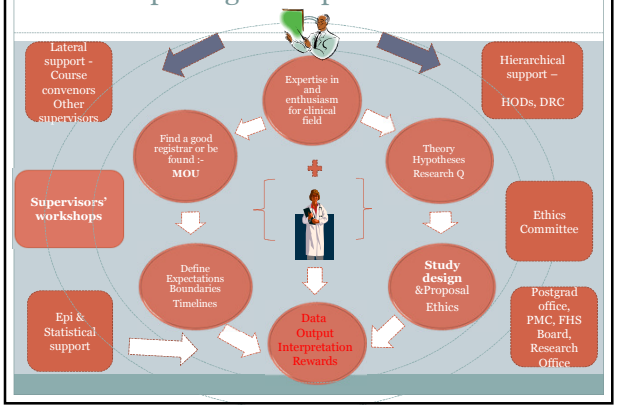
Supervisor

- Pass on knowledge
- Guidance/direction, not counselling
- Oversee the process A-Z
- Rules, Requirements, Forms, Approvals, Ethics
- Provide contacts/access
- Thesis format choice
- Control funding

Problems with transitioning

- May have only done some of the above
- Very little previous experience
- May feel coerced
- May not be assigned a research-ready registrar
- Need to see the benefits to your own career
- Need a co-supervisor, if available
- Need departmental and Faculty support

Exploring the supervisor's role



Support for supervisors / PG students

- Handbooks and links
- VULA: Resources:- Supervisors handbook; all PG forms; the PG office can register you onto the PG VULA site.
- Workshops –
- Registrars Research Methods course: 24/31 May 2014
- Study design/epi support/sample size and stats
- http://medstats.uct.ac.za/
- Ethics advice: Marc Blockman
- http://www.health.uct.ac.za/research/humanethics/about/
- Mentoring sessions
- Follow-up to this WS, Department specific needs, by arrangement

Benefits to you

- Improve your CV - supervision of successful candidates is a credit to you
- Pass on hands-on aspects of the study, frees up your time
- Publication potential increases
- Attracts some funding into your department
- Increases your capacity as a researcher and a leader/supervisor
- Improves your ability to become a Principal Investigator

Student vs supervisor

- It is the responsibility of the student to conform to University and programme requirements and procedures with regard to such matters as research ethics, dissertation style etcetera.
- Although it is the duty of the supervisor to be reasonably available for consultation, the primary responsibility for keeping in touch rests with the student.

Form D2:Memorandum of understanding

University of Cape Town
Faculty of Health Sciences

Between postgraduate student and supervisor
This memorandum of understanding between
.....(name of graduate student)
and
.....(name of supervisor)
is designed to ensure that the supervision experience is as mutually productive as possible.

Outline of expectations and commitments:
Research expectations: *(laboratory access; field work; access to equipment; courses to attend; conference attendance; seminar presentations)*

Supervisor/student commitments *(access to supervisor; annual leave for student; working hours; (co)authorship of articles)*

Financial support: *(stipend; research costs; conference and travel, etc):*

University of Cape Town
Faculty of Health Sciences

Memorandum of understanding between postgraduate student and supervisor

This memorandum of understanding between

……..(name of graduate student)

……..…….(signature)

……(date)

……(email address)

……..(cellular phone number)

and

……..……..……..……(name of supervisor)

……(signature)

……...…….(date)

…….(email address)

……..(contact number)

is designed to ensure that the supervision experience is as mutually productive as possible.

Further, it has been discussed and noted by

……..……..……..……..……(name/s of co-supervisor/s)

……(signature/s)

……...…….(date)

and

……..……..……..……(name of co-supervisor)

……(signature, if available)

……...…….(date)

(This must be completed within six months of initial registration; an annual **'progress and planned activity'** report must be completed each subsequent year before the student renews his/her registration. Signatures on the submitted form must be original, other than if a co-supervisor is outside of UCT and thus not easily available. In such a case the co-supervisor should indicate by way of written/recordable means (e.g. email) that they have seen this MoU and are agreeable to serve as co-supervisor. This can be attached to this MoU.)

Candidate details:

A1 Name of Candidate: _____ Student number: _____

A2 Highest academic qualification: _____

A3 Degree registered for: M ____ PhD ____Year of first registration: _____

A4 Project title and proposal: (attach proposal separately):

……

……

The supervision arrangements:

General obligations of supervisors are outlined in appendix I. By signing this document, both parties acknowledge their understanding of the general expectations it contains.

B1 Supervisor:

(a) Initials & surname: ……

(b) Staff no:..……..……

(c) Department:…..……

B2 Co-supervisor(s) if any:

(a) Initials & surname:…….

Department:……..……..

Email: ……

Institution:…...……

Responsibilities:…...……..

……

……

……

(b) Initials & surname:……...……..

Department:……..……..

Email: ……

Institution:…….…….…….……..

Responsibilities:……...…….

……...……

……

……

Outline of expectations and commitments:

C1 In free format/point form provide an outline of expectations set out in as much detail as possible to the satisfaction of supervisor and candidate (use separate pages if necessary).

Research expectations: *(laboratory access; field work; access to equipment; courses to attend; conference attendance; seminar presentations)*

……

……

……

……

……

……

……

……

……

……

……

……

……

……

Supervisor/student commitments (*access to supervisor; annual leave for student; working hours; (co)authorship of articles):*

……..

……..

……..

……..

……..

……..

……..

Financial support: *(stipend; research costs; conference and travel, etc):*

……..

……..

……..

……

……..

Intellectual Property

1. As the student, by signing this document, I confirm that I have read the UCT IP Policy (www.uct.ac.za/about/policies/).
2. Who funds the research (exclude bursaries)? _____
3. In terms of the funding arrangement, has the IP been assigned to the funder (i.e. either because the full cost model has been applied to the project, or in terms of a research contract)? **YES / NO** (delete the non applicable)
4. In terms of the IP Rights from Publicly Financed Research and Development Act, the Student and Supervisor acknowledge that in all cases where the answer to 3 is "No" there is an **obligation to disclose** an invention to Research Contracts and IP Services with 90 days of the discovery, using an Invention Disclosure Form (download from www.rcips.uct.ac.za/ip/overview/). There is an obligation to maintain the invention confidential within UCT until the IP has been evaluated by RCIPS to determine its ability to be protected. RCIPS should be contacted well in advance of any planned public disclosure, such as presentation at an external meeting or conference, publication in a journal, submission of an abstract, publication on a website or blog and the submission of a thesis for examination.
5. In terms of the UCT IP Policy, the university owns the IP arising from postgraduate research (except for copyright in a thesis, as per Clause 6) unless ownership has been assigned to a third party. This includes inventions, discoveries and other developments of a technical nature whether or not these may be the subject of legal protection, as well as tangible research property arising from research activities such as prototypes, drawings, designs and diagrams, biological organisms and material, reagents, integrated circuit chips, software and data.
6. Copyright in a dissertation or thesis vests in the student who has written the dissertation or thesis, subject to the usage rights of the University provided in rules for degrees, diplomas and certificates. In terms of Rule GP8, when presenting a thesis for examination, a candidate shall be deemed by so doing to grant a royalty-free, non-exclusive, non-transferable licence to the University to publish it in whole or in part in any format that the University deems fit. Students should take note of this provision should they enter into an agreement with a publisher to publish their thesis.
7. The University assigns the copyright of all scholarly and literary publications to the authors of such works.
8. Graduate students often use data that belongs to the University, or a research group, or an external party. Any issues relating to data ownership should be noted here: _____

_____

Observation by Head of Department or (if appropriate) Head of Division

I have reviewed this completed MoU and I am satisfied that the department and division (if applicable) is able to meet the obligations to the candidate as set out in this MoU:

Signed:…….…… (Head of Department)

Name:..……..……

Date:..……

Signed:…….…… (Head of Division)

Name:..……..……

Date:..……

I approve registration of the candidate in the Faculty of Health Sciences:

Signed:……

Dean/Dean's nominee

Name:..……

Date:……

POSTGRADUATE OFFICE

MANAGER: ADRI WINCKLER
REGISTERED STUDENTS:
LORRAINE MC DONALD & DIANNE PRYCE
REGISTRARS: THOBEKA MNGAZA

DEADLINES

- Registration: Last **Friday** in February and must register for the **full coursework component** at that time. Penalties occur for late registration!
- When the candidate starts preparing for the research component for dissertation, he/she should contact the PG Office to register for the dissertation component.
- This should be when ethics approval letter has been received. The PG office sends all forms to PMC and Faculty Board for approval, then the study can start.
- Registrars (MMed students) and subspeciality trainees (MPhil students) are required to register annually with the Health Professions Council of South Africa.

COMMUNICATION

- **Student email:** UCT address. If using different email they must use UCT email at least once to set up an auto-forward to their preferred email address.
- **VULA** has a postgrad site that all are automatically registered on. Handbooks, tutorials, lecture notes, venues/times, **scholarships, bursaries, and forms**
- **SMS:** Correct cellphone numbers must be given to postgrad office, especially new/changed numbers

DISSERTATION REQUIREMENTS

- A 60 credit dissertation should be no more than 20,000 words in length eg. MMed.
- The dissertation of 90 credits of a coursework master's degree should be no more than 25,000 words in length
- A degree by full dissertation should be no more than 50,000 words in length.

DISSERTATION FORMAT – D4

Publication ready	Monograph Format
Part A: The REC approved protocol (4000w) **Part B:** Structured literature review (3-4000w) **Part C:** Publication ready manuscript (3000w) for a named journal following instructions to Authors **Part D:** Appendices	16 000 to 20 000w in a comprehensive and scholarly style **A:** Structured literature review **B:** Materials and methods, (appendices) **C:** Results, discussion and conclusions

SUBMISSION

- Notice of intention to submit a dissertation shall be given in writing to the Faculty Office, no later than 1 month before submission, **Form D8**.
- Not later than 15 March for June graduation, and 15 August for December graduation.
- ***NB: 100% fee rebate if dissertation submitted on 1st day of academic year (feb), or 50% rebate before start of 2nd Semester (mid Jul)***
- No of copies:
  - 2 copies in temporary binding
  - 1 CD of the dissertation in one continuous file in a universally readable format (i.e. Word or PDF)

EXAMINERS

- Candidates should submit an abstract (500 words) of the research project no later than 6 months before submission.
- Supervisors should send full contact details for 3 **external** (non-UCT) examiners to be invited on nomination Form from PG Unit
- *Those invited may not be appointed, therefore be diplomatic with the invitation.*
- Supervisor must provide full detail of nominated examiners, including highest qualification obtained and description of their standing in the relevant field of research
- Must include complete information, namely:
  - Delivery address with postal code (NO postal addresses)
  - Fax number (incl. ALL codes) and Work tel. no (incl ALL codes)
  - Most recent email address

The candidate should not be informed of the identity of the examiners.

EXAMINATION BOARDS

- Examination results must be captured on PeopleSoft as soon as possible by the department admin but at least 3 days prior to the relevant exam board meeting
- A signed (by HOD) hard copy of the results must be sent to the PG Unit by admin (for record purposes)
- PG Unit will print CRS (course results schedules) for use at the exam board meetings

FORMS TO BE COMPLETED

- **Form D2a**: MOU-Memorandum of understanding. Contract between you and your supervisors. Should be filled in within 6 months of registration for an MMed.
- **Form D3**: Appointment of supervisors.
- **Form R1:** Submit proposal of research topic to Departmental Research Committee (DRC) with for approval in the format for eventual submission to Ethics Committee. (4000 words max).
- **FHS013 (DRC) and FHS015**: Submit proposal to Ethics committee (forms on website).
- **Form D1**: Submission of study proposal after REC approval plus
- *These forms, **D1, D2a and D3** should be submitted to the Post Graduate Office once ethics approval is granted with ethics approval letter and study proposal. PG sends to FHS board.*
- **Form D8**: Intention to submit at least one month before submitting for examinations.

ETHICS REQUIREMENTS

PROCESS AND FORMS

- After DRC approval of protocol on form R1 Submit cover letter to Ethics, Room 24, E52, Old Main Building, Groote Schuur Hospital
- FHS013 – application for REC approval
- FHS014 – synopsis (2 pages)
- FHS015 – Protocol and ethical considerations
- C1online (FHS002) – if budget is to go through FHS
- FHS020 – Application to register a database

SUBMISSION TO FHS

- The full protocol FHS015
- Form D1
- REC approval letter
- Must all be submitted to the postgrad admin office for approval by the *Professional Masters Committee chair* and the *Board of the Faculty of Health Sciences* before starting the research
- MMeds should aim for completion by end of year 2, or 6 months before writing the Part II exam

ETHICS APPROVAL

- Primary research taking place in provincial or local authority health facilities must be submitted to provincial Govt. for approval after UCT REC
- However, teaching hospitals and local authorities approve research projects in-house for access to public sector facilities
- 'All other province' approvals by the Directorate: Health Impact Assessment (Research) at provincial head office. (Prof. Rodney Ehrlich is chair of PHRC)

FORM FHS015: RESEARCH PROTOCOL – SECTION C

The protocol must reflect how the research will be conducted at the local research site, for example contact details of the local principal investigator (PI) and Human Research Ethics Committee, characteristics of the local population and information about recruitment sites.

Instructions

- Forms to be downloaded from the Administrative Forms web page at
- http://web.uct.ac.za/depts/sapweb/forms/forms.htm
- All researchers must complete Section C

Good research conduct is very important:
Good Clinical/Laboratory Practice training

SECTIONS TO COMPLETE ON FHS-015

1. **Purpose of the study**
2. **Background**
3. **Methodology**
4. Study design
5. Characteristics of the study population
6. Recruitment and enrolment
7. Research procedures and data collection methods
8. Data safety and monitoring
9. Data analysis
10. ***Description of risks and benefits***
11. ***Informed consent process***
12. ***Privacy and confidentiality***
13. **Reimbursement for participation**
14. **Emergency care and insurance for research-related injuries**
15. **What happens at the end of a study?**
16. **References and Appendices**

SUMMARY

Where	What	Forms/ time
PG Office	Course registration	Last Friday in Feb
PG Office	Dissertation registration	After ethics granted
DRC	Study proposal *	D1, FHS013, C1
Ethics Committee	Covering letter, proposal, synopsis, budget Database registration	FHS013 FHS015, FHS014 FHS002/C1 FHS020
PG Office	Ethics approval letter, protocol MOU Supervisor apptmt	D1, D2a within 6mo of reg D3
PMC, FHS Board	Approval of study	Study start *
PG Office	Intention to submit	D8

Summary

Where	What	Forms/ time
PG Office	Course registration	Last Friday in Feb
PG Office	Dissertation registration	After ethics granted
DRC	Study proposal *	D1, FHS013, C1
Ethics Committee	Covering letter, proposal, Synopsis, budget Database registration	FHS013 FHS014, FHS015 FHS002/C1 FHS020
PG Office	Ethics approval letter, Protocol MOU Supervisor appointment	D1 D2a within 6mnths of reg D3
PMC, FHS Board	Approval of study	Study start *
PG Office	Intention to submit	D8

ROLES AND RESPONSIBILITIES OF THE SUPERVISOR AND THE GRADUATE

The responsibilities of the supervisors:

- Guiding and advising the student on the selection and development of a research topic that is challenging, at the appropriate level for the degree sought and can be completed within the expected time frame of the degree program. The supervisor should ensure that the research receives approval from the appropriate Research Ethics Board if required prior to commencement of the project. Supervisors should be mindful of the availability of the resources needed to pursue the research.

- Communicating to the student the required levels of performance, as well as the performance indicators that are consistent with satisfactory and timely progress in the degree program.

- Ensuring that the student gains the necessary theoretical foundations and acquires the skills required to conduct research in a manner consistent with the highest standard of ethical and scientific practice.

- Examining thoroughly and responding in a timely manner (usually within 2 weeks) to written work relevant to the thesis/research project submitted by the student and providing constructive suggestions, preferably in writing, for improving and continuing the work.

- Informing the student of the standards for quality and style to which theses and papers for publication must conform and advise their students accordingly.

- Being accessible to the student to consult and discuss their progress. The frequency of such meetings depends on the field of study, the type of program, the stage of the research project and the independence of the student. At a minimum, meetings should be arranged in each academic term.

- Ensuring a safe research environment compliant with University and departmental regulations that is supportive of the research enterprise and is free from discrimination, intimidation or harassment.

- Making arrangements of continuity of supervision during periods of extended leave or absence (greater than 6 weeks) in consultation with the student and informing the graduate coordinator.

- Advising the student on career options and opportunities for professional development.

- Encouraging and assisting the student to disseminate the research findings through appropriate channels (conference, meetings, journals, etc.)

- Being informed of the regulations and procedures of their department, the School of Graduate Studies and the University and being knowledgeable about services and resources available to graduate students at UCT.

- Being honest and open with the student about expectations, performance, and all issues relevant to the student's academic progress.

Roles and responsibilities of the graduate students:

When a student registers in a graduate program, the student makes a commitment to strive for the highest levels of academic achievement and contribute fully to the intellectual life of the University. The primary responsibility of the student is to devote the time and energy required to complete all academic requirements, including the thesis within the expected time frame. It is also the responsibility of the graduate student to follow departmental, School of Graduate Studies and University policies, procedures and regulations and to adhere to the principles of academic integrity. By agreeing to work with a supervisor, students enter a partnership that will succeed if it is built on mutual trust and respect. Students should acknowledge the senior role that is played by their supervisors who are experienced researchers and academics. It is expected that students should seek their supervisor's advice and give it serious consideration. Students should also seek advice and guidance from their supervisory committee members and from their Graduate Coordinator as needed.

<u>The responsibilities of graduate students include:</u>

- Becoming familiar with and complying with the policies, procedures and regulations of the department, Graduate School OF Studies and the university.

- Meeting deadlines and following regulations associated with registration, award applications, and degree requirements including thesis format and submission.

- Conducting research with the highest standard of ethical and scientific practice and acquiring Research Ethics Board approval if required.

- Providing reasonable evidence of satisfactory research progress, as requested by the supervisor or supervisory committee.

- Preparing a research plan and timeline in consultation with the supervisor as a basis for monitoring progress and completing all stages of the research.

- Giving serious considerations to the advice and criticisms offered by the supervisor and the supervisory committee regarding academic and research work.

- Keeping the supervisor informed of progress and research findings through regular meetings and open communication.

- Ensuring that contact information is up to date with the supervisor, the department, and the registrar's office.

- Informing the supervisor of any changes that might affect progress.

- Seeking advice and support from university services and resources as needed.

- Upon completion of the research work ensure that all records, files, and documents are stored appropriately and a plan for dissemination has been agreed upon by all collaborators.

Useful Links

Useful websites/articles/contacts

Research methodology and statistics

RD Direct – Research Flowchart
http://www/rdinfo.org.uk/flowchart/Flowchart.html

Survey and Questionnaire Design Resources
http://www.statpac.com/surveys/

Tom Lang. Twenty Statistical Errors Even YOU can Find in Biomedical Research Articles. Croatian Medical Journal. 45(4):361-370,2004

Reporting guidelines for manuscripts

http://www.equator-network.org/resource-centre/library-of-health-research-reporting/reporting-guidlines/

Reporting guidelines are statements that provide advice on how to report research methods and findings. Usually in the form of a checklist, flow diagram or explicit text, they specify a minimum set of items required for a clear and transparent account of what was done and what was found in a research study, reflecting in particular issues that might introduce bias into the research.
Most widely recognised guidelines are based on the available evidence and reflect consensus opinion of experts in a particular field, including research methodologists and journal editors. Reporting guidelines complement advice on scientific writing, which concentrates on the basic writing principles and styles of research reports and publications, and journals' instructions to authors.

Initiatives to improve research design, conduct and reporting

Core outcome networks /groups /collaborations
http://www.comet-initiative.org/links

MIBBI – Minimum Information for Biological and Biomedical Investigations
http://mibbi.sourceforge.net/

Blogs, online discussions, collections on reporting

BiomedCentral – Open Data
http://blogs.openaccesscentral.com/blogs/bmcblog/category/Open+Data

BMJ Research Methods & Reporting
Collection of published papers

PLoS Medicine Guidelines and Guidance
Collection of published papers

Editorial organisations and guidelines

ICMJE – International Committee of Medical Journal Editors: Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication
http://www.ecmje.org/

Wager E. Getting research published: an A to Z of publication strategy. Radcliffe Publishing Ltd, Second Edition, 2010

Hall GE 9ed). How to write a paper. Fourth Edition. MBJ Publishing Group, Wiley – Blackwell, 2008.

Byrne D. W. Publishing your medical research paper: What they don't teach in medical school. Williams & Watkins, Baltimore, US, 1998

Publication Manual of the American Psychological Association, Sixth Edition. APA Washington DC, US, 2009

Fraser J., Fuller L. & Hutber G. Creating Effective Conference Abstracts and Posters in Biomedicine. Radcliffe Publishing, UK, 2009

Lang, T. How to Write, Publish & Present In The Health Sciences: A Guide for Clinicians and Laboratory Researchers. Philadelphia, PA: American College of Physicians, 2009

Writing a Thesis

Helpful sites for thesis writers seem to be proliferating but this article provides a selected few based on extensive evaluation.

http://www.woodhillpark.com/blogs/25/Helpful-sites-for-thesis-writers-seem-to-be-proliferatimg.html

http://www.timeshighereducation.co.uk/news/how-not-to-write-phd-style-thesis/410208.article

Professional organisations and associations

International Network for the Availability of Scientific Publications (INSAP)
http://www.insap.info/

International Society for Medical Publication Professionals (ISMPP)
http://www.ismpp.org/

Copyright Information

Creative Commons
http://creativecommons.org

UK Copyright Service
http://www.copyrightservice.co.uk

US Copyright Office
http://www.copyright.gov/

Other useful links

Cochrane Collaboration
http://cochrane.org

Instructions to Authors (Mulford Library, University of Toledo)
http://mulford.meduohio.edu/instr/

James Lind Library – helping to understand fair tests of treatments in health care
http://www.jameslindlibrary.org

MORE – McMaster Online Rating of Evidence
http://hiru.mcmaster.ca/More/AboutMore.htm

UCT Support

Free research support service from School of Public Health & Family Medicine

- Email: pph-epsupprt@uct.ac.za

Systematic reviews

- James Irlam, James.Irlam@uct.ac.za

Pulmonology/Occupational

- Rodney Ehrlich, Rodney.Ehrlich@uct.ac.za

HIV; Obs/Gynae; Ophthalmology

- Landon Myer, Landon.Myer@uct.ac.za

Critical Appraisal

These checklists are extremely helpful for 2 reasons:

1. They train research design
2. They provide the student with the necessary information to be able to critique an article in an informed manner.

http://www.sign.ac.uk/methodology/checklists.html

http://www.casp-uk.net/

(A) Case-control studies

Definition of study groups:

- Describe an objective definition for the disease or outcome of interest and use this to define the **cases**.
- Describe the appropriate control group for the particular group of cases. In general the controls should be sources from a similar population as the cases, the only difference being that they do not have the outcome of interest.

Matching (if relevant):

- Describe the matching factor(s) and how the matching will be done. The main purpose of matching is to control for confounding. By matching on age, for example, we are attempting to ensure the cases and controls have the same (or similar) age distribution.

Blinding and avoiding bias:

- List the measurements that are to be blinded and how this will be done. Blinding is a technique used to minimise bias In observational studies, the minimisation of ascertainment bias can be helped by blinding. e.g. If subjects in a case control study are to be interviewed to assess exposure status, then the interviewer should, if possible, be blind to the outcome of the subjects.
- Methods used to minimise other types of bias should be described. E.g. using multiple sources of information can help minimise verification bias.

(B) Cohort Studies

Definition of study groups:

- The exposure of interest should be clearly described and used to define the "exposed" group. Groups like doctors, civil servants, surgery patients are often chosen as the source of the groups because they are easy to define or monitor. In these situations an internal comparison group without the exposure of interest would be used. E.g. doctors who smoke and doctors who do not smoke. Sometimes groups with special exposures are chosen (e.g. brewery workers, workers in nuclear power stations) and in these cases an external comparison group would need to be found.

Blinding and avoiding bias

- List the measurements that are to be blinded and how this will be done. Blinding is a technique used to minimise bias. E.g. in a prospective cohort study where subjects are being followed up in the future, the researcher who assess any outcome measures should, if possible, be blinded to exposure status.

(C) Cross-sectional studies

In cross-sectional studies the exposure status and disease status are assessed at a single point in time (e.g. a sample survey; 2001 census). It is therefore **not possible** to establish whether the exposure preceded or resulted from the disease. However, data from these studies are useful in providing information about the health status and needs of a population. They are sometimes called prevalence studies. Selecting the study groups has the same problems/ features as with cohort studies. Response rates are also often a problem and the strategy used to recruit a representative sample and for maximising the response rate should be given.

(D) Other types of studies

Other types of observational studies may involve only one group of subjects. For example, in a method comparison study, measurements will be taken on a group of subjects using two (or more) methods. The guidelines set out in this document still apply, the main difference being that a second group of subjects is not used.

9. Study group(s)

Include details of:

- Source of subjects (where they come from and why this group is appropriate). For example, are they a random sample from a larger population or are they obtained from a disease register? They may be patients attending a clinic over a particular period.
- Number of centres involved or regional / national limits on possible recruitment
- Subject inclusion and exclusion criteria (with justification if necessary)
- Expected no of eligible participants available per year and proportion of these expected to agree to take part.

10. Recruitment or choosing participants

Details of recruitment process including

- method of recruitment (e.g. via adverts, clinics, GP referral, entry in established database (e.g. Society of Cardiothoracic Surgeons National Database)
- payment of participants
- details of procedures, tests, screenings carried out to assess study suitability
- Provision of patient information sheet (include as appendix)
- gaining patient consent (how consent will be obtained, who will gain consent, whether a witness will be present, how long the subject will have to decide, the arrangements for non English speakers and special groups (e.g. mentally ill, children, those suffering from dementia.)
- detail of enrolment procedure

11. Data

11.1 Data to be collected

- if using an established database permission to use the data must be sought and a detailed description of how (and by whom) the data will be collected for the study should be given. Also, an example of the data extraction form used to obtain the data for your study should be included.
- provide a detailed list of all data (outcome variables, explanatory variables, potential confounding variables etc) to be collected, with each description including :
  - source of the data (e.g. patient questionnaires, patient notes, electronic data, procedure)
  - time point for collection (baseline, during treatment, at follow up point)
  - who will collect the data
  - why the data is being collected (e.g. baseline comparison data, primary outcome, important prognostic / explanatory /confounding variable)
  - whether the data is from a standardised tool (e.g. McGill pain score) involves a procedure (in which case full details should be supplied). If a non-standard tool is to be used, detail on reliability and validity should be given.
  - what form the data will take (e.g. binary, continuous (numeric), time to event)

- useful to include table / diagram describing schedule for data collection.
- describe methods used to maximise completeness of data (e.g. telephoning patients who have not returned postal questionnaires)
- include data collection forms and questionnaires as appendices

11.2 Data handling and record keeping

- describe procedures for data collection and recording (software to be used, location of the data etc)
- detail methods implemented to ensure validity and quality of data (e.g. double entry, cross validation etc)
- Security / storage of data
- Records retention – duration and location
- Adherence to Data Protection Act 1998 and Caldicott

12. Statistical Considerations

12.1 Sample size calculation

Details of the precision or power calculation used to estimate the required sample size based on primary outcome

- Assumptions made (statistical assumptions regarding distribution)
- Estimates of difference to be detected along with appropriate justification
- Chosen levels of significance and power.
- Reference or details of method/formula used for the calculation

12.2 Analysis

- Describe which variables will be used to assess groups comparability and how they will be reported (e.g. means, proportions)
- Description of primary and secondary analyses including summary measures used, methods of analysis (e.g. t-test , logistic regression) and how the results will be reported (e.g. odds ratios with 95% confidence intervals)
- Details of adjustments for pre – defined confounders.
- Approach used to deal with missing data and loss to follow-up.
- When will the analysis be done and by whom.
    - Details of how any planned subgroup analyses will be done

13. Compliance

13.1 Subject compliance

- procedures for monitoring (e.g. exercise diary for subjects on a rehabilitation programme)
- recording of patient compliance information (what will be recorded and where)
- detail of follow-up of non compliant subjects

13.2 Withdrawal of subjects

- describe under what circumstances and how subjects will be withdrawn from the study
- give details of documentation to be completed on subject withdrawal (including recording reasons for withdrawal and any follow-up information collected)

14. Ethical Considerations

Description of ethical issues related to the study. For example consider:

- Approvals from relevant groups (e.g. MREC, LREC, MHRA, Trust(s))
- Informed consent (subject information and informed consent form appended)
- Allowances for special groups (e.g. non English speakers, children, mentally ill)
- Patient withdrawal / discontinuation

15. Finance and Insurance

- Finance and insurance details (if not addressed in separate agreement)
- Cover for non negligent and negligent harm

16. Reporting and Dissemination

Details of how, where and when the results of the study will be reported / presented.

Useful reading

Websites

- Declaration of Helsinki (http://www.wma.net/en/30publications/10policies/b3/index.html)
  Provides ethical principles for medical research involving human subjects
- COREC guidelines (www.corec.org.uk)
  Includes patient information sheet and consent form guidelines
- Martin Bland et al, Statistical Guide for Research Grant Applications,
  http://www-users.york.ac.uk/~mb55/guide/guide.htm
  Includes detailed information and definitions of many aspects required for a research protocol as well as information about randomisation software and services
- ICH Harmonised Tripartite Guidelines for Good Clinical Practice
  (http://www.cgmh.org.tw/intr/intr1/c0040/web/C/ICH%20GCP%20E6.pdf)

Books

- Altman, DG. (1991) *Practical Statistics for Medical Research*. London: Chapman and Hall, Chapter 15.
- Bland M, (1995) *An Introduction to Medical Statistics*, Oxford University Press
- Betty R Kirkwood and Jonathan AC Sterne (Blackwell *0-86542-871-9)*
- *Ed Douglas Altman, David Machin, Trevor N Bryant, Martin J Gardner (BMJ Books ISBN 0 7279 1375 1) Statistics with Confidence*
- Rothman K.J, Greenland S, *Modern Epidemiology* (2nd Edition). Lippincot-Raven
- Hennekins C.H., Buring J.E.(1987) *Epidemiology in Medicine*. Lippincot-Raven (excellent chapters on design and bias)
- Machin D, Campbell MJ, Fayers PM and Pinol APY (1997) *Sample Size Tables for Clinical Studies.* Second Edition. Oxford: Blackwell Science.

Papers

BMJ statistics notes provide some brief but useful information on various topics including:

- J Martin Bland & Douglas G Altman, Matching BMJ 1994;309:1128 (29 October)
- J Martin Bland & Douglas G Altman Multiple significance tests: the Bonferroni method BMJ 1995; 310:170
- J Martin Bland & Douglas G Altman *Presentation of numerical data* BMJ 1996; 312:572 (2 March)
- J Martin Bland & Douglas G Altman Survival Probabilities (the Kaplan-Meir method) BMJ 1998; 317:1572-1580
- J Martin Bland & Douglas G Altman The logrank test BMJ 2007; 328:1073
- J Martin Bland & Douglas G Altman Analysis of continuous data from small sample BMJ 2009; 338:a3166
- DG Altman, JM Bland BMJ One and two sided tests of significance 1995; 311:485
- D G Altman, JM Bland Comparing several groups using analysis of variance BMJ 1996; 312:1472-1473
- DG Altman, JM Bland Bland Units of analysis BMJ 1997; 314: 1874
- DG Altman, JM Bland Time to event(survival) data BMJ 1998; 317:468-469
- DG Altman, JM Bland Standard deviation and standard errors BMJ 2005; 331:903
- Day JD, Altman DG. *Blinding in clinical trials and other studies BMJ* 2000; 321: 504
- David A Grimes, Kenneth F Schulz Descriptive studies what they can and cannot do Lancet 2002; 359:145-49
- David A Grimes, Kenneth F Schluz Bias and causal association in observational research Lancet 2002; 359:248-52
- David A Grimes, Kenneth F Schulz Case-control studies: research in reverse Lancet 2002; 359:431:34
- David A Grimes, Kenneth F Schulz Compare to what? Finding controls for case-control Lancet 2005; 365:1429-33

Other useful articles

Statistics Notes in the ***British Medical Journal***

This series is edited by Doug Altman, Cancer Research UK, and Martin Bland, University of York. All are available on the BMJ website.

For an extended list of publications on statistics and research methods use the BMJ link to topic collections www.bmj.com/cgi/collection and go to non-clinical - statistics and research methods.

The Research Process

The Clinical Research Process

CELESTE DE JAGER
SCHOOL OF PUBLIC HEALTH
FHS
SUPERVISORS WORKSHOP 2

What is required

- Independent work
- Evidence of ability to undertake research
- Evidence of ability to interpret results adequately
- Ability to review literature comprehensively
- Ability to review literature critically
- Publication-ready or monograph format

Overview of process

1. Registration and orientation
2. Supervisor/Topic choice/Lit review
3. Basic protocol (4000w) with supervisor's help.
4. *Meeting to discuss study protocol (CdJ*
5. *Meeting to discuss sample size calcs, data entry and stats plan: (Henri/Celeste)*
6. Protocol submission for internal review (DRC)
7. **FHS-013** and **FHS015**, then to ethics
8. Prep for research period: data file set-up, folder/patient /lab access

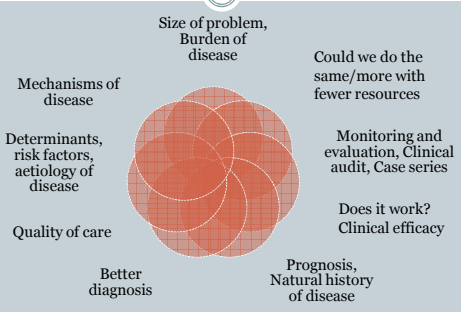
Overview of process

8. Data collection process
9. Departmental progress updates for rotating registrars; presentations/journal clubs
10. *Analysis: help from supervisor/meeting with Henri/Celeste/other*
11. Discuss interpretation with statistician/supervisor
12. Submission-notify postgrad office in time- **FormD8**
13. Possible conference/ FHS Research Day
14. Possible publication

Start with a good study question

- **Relevant**
  - Addresses topic of significance to health of local population / health care services
- **Novel**
  - Makes meaningful contribution to existing knowledge ~ new insights
- **Ethical**
  - Does not interfere with normal standard of care or put people at risk of unnecessary or unknown harm
- **Feasible**
  - Not overly ambitious!!

Topic Choices



Specific

- Who? population or cases
- What? Factor/exposure /intervention/comparison
- Why? Outcome/endpoint/association
- Size/impact and direction of effect :- clinical significance versus statistical significance

What makes a research question scientific?

- Has a structured method that is transparent
- Can be interpreted in a transparent way
- Is replicable by others
- Can be distinguished from popular opinion or experience i.e. evidence-based
- Should yield information and knowledge beyond the original situation
- But can include trial and error/experimental intervention
- Capable of withstanding peer review

Refining the question

- Have a general idea of the problem
- Do literature review of recent articles/discuss
- Formulate the justification/value of research
- Draft the hypothesis/question/aim
- Convert into practical objectives
- Ascertain if the question has been previously answered by others
- Consider the ethical and feasibility aspects
- Refine question and formulate the methods involved

Summary of question criteria

- **F**easible
- **I**nteresting
- **N**ovel
- **C**ontrary
- **E**thical
- **R**elevant/justifiable
- **S**pecific

Note: *You may have a study appropriate for subdivision into a number of smaller studies for MMeds, or add-on projects.*

Research question

- Do a retrospective review of all patients seen with X disease from 19xx to 20xx and look at the demographics, aetiology, risk factors, co-morbidities, severity, management, and survival, compare with other studies done and make recommendations for the local handbook ….” OR

- ‘Determine the incidence of hip fractures in old age homes in Cape Town in a 1-year period’

Key issues

1. Simplify topics and study designs

2. Turning Aims into Objectives –

Means stating how the aims will be achieved; what measures will be used to determine the answers to the questions posed.

- Don’t include extra objectives over and above aim. Unexpected findings or extra data can be presented, but the examiner will look for answers to the objectives, so don’t include more than 2 or 3 objectives.

3. Early stats/design advice

4. Consider data entry options

Example

- Chest X-rays are expensive. Is it necessary for every child with a lower respiratory infection presenting to clinic to have one?
- **Hypothesis:** Children with LRTI seen at RXH OPD who have a chest x-ray taken have the same time to recovery as children who do not have an x-ray.
- **Objective:** To measure time to recovery in children with and without chest x-ray for LRTI at RXH OPD.
- ***Secondary objective:*** To measure the relapse rate between the two groups

Exercise

- Depression after stroke is common and difficult to manage for the health services and family. How common is it in fact, and can we identify some modifiable factors that contribute to post-stroke depression?
- State the justification, hypothesis and objectives for a research study on this topic.
- Identify the population, exposures/factors and outcomes/measures that will be included in the study.

Options in study design

- – Case report  Not acceptable for MMed
- – Case series  eg. retrospective/XS
- – Case - Control comparison  vs. 
- – Cohort  Longitudinal
- - Lab study: tissue, blood, other body fluids, animals

Study designs

Observational

- Descriptive
- Analytical
- Experimental

Prevalence
Incidence
Demographics

Risk factors
Exposure x Outcome
Diagnostic

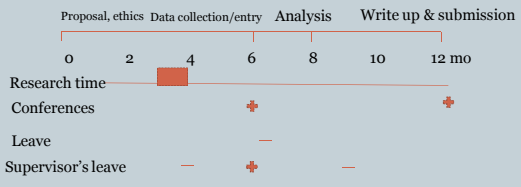
Community study
Clinical trial, RCT
with intervention

Study Outcome measures and Data entry

Outcomes	0mo	6mo	12mo
Med history	X		
Medications	X	X	X
Disease screen	X		
Inclusion criteria	X		
Questionnaires	X	X	X
Blood tests	X		X
BP, Heart rate, vibration sensitivity	X		X
Treatment	X	X	X
Treatment outcome	X	X	X

Project timetable and milestones

During the 4 year MMed, the research project should be completed in good time



Form FHS015: Research Protocol – Section C

The protocol must reflect how the research will be conducted at the local research site, for example contact details of the local principal investigator (PI) and Human Research Ethics Committee, characteristics of the local population and information about recruitment sites.

Instructions

- Forms to be downloaded from the Administrative Forms web page at
- http://web.uct.ac.za/depts/sapweb/forms/forms.htm
- All researchers must complete Section C

Good research conduct is very important: Good Clinical/Laboratory Practice training

Sections to complete on FHS-015

1. **Purpose of the study**
2. **Background**
3. **Methodology**
4. Study design
5. Characteristics of the study population, (incl sample size calc.)
6. Recruitment and enrolment
7. Research procedures and data collection methods
8. Data safety and monitoring
9. Data analysis
10. ***Description of risks and benefits***
11. ***Informed consent process***
12. ***Privacy and confidentiality***
13. **Reimbursement for participation**
14. **Emergency care and insurance for research-related injuries**
15. **What happens at the end of a study?**
16. **References and Appendices**

Statistical Support

UCT CLINICAL RESEARCH CENTRE
every step of the study

What the CRC does with MMeds

STATISTICAL SUPPORT

@ MMed Research Training Days

@MMed Focused Sessions

@ Individual consultations

What the supervisor could focus on

STATISTICAL SUPPORT

@ MMed Research Training Days

- Converting an idea into a question
- Study designs
- Selection and sampling
- Quality control of measurement
- Summarizing different data types /database design
- Introduction to sample size calculation
- Ethical issues
- The research protocol and final write up

STATISTICAL SUPPORT

@ MMed Focused Sessions

- Mmed supervision and Process
- Critical appraisal
- Database searches
- Calculating sample size
- Prevalence, Odds ratio, Risk ratio
- Agreement, Sensitivity, Specificity
- Association /Correlation

UCT CLINICAL RESEARCH CENTRE

STATISTICAL SUPPORT

@ Individual consultations

- Answer questions specific to the analysis of the particular project (e.g. which test is more appropriate)
- Refer to Biostatistician if necessary

UCT CLINICAL RESEARCH CENTRE

STATISTICAL SUPPORT

What the supervisor could focus on doing with the student

- Go over the aims and specific objectives
- Translate this into an analysis plan (broadly)
- Advise on dummy tables for the results
- Assist student with interpretation of the results

UCT CLINICAL RESEARCH CENTRE
every step of the study

Research Question		Design possibilities
1	**What is?** **What are?** *What is the prevalence of fear of falling among community-dwelling older persons with impaired mobility*	Survey Exploratory Descriptive Case study Needs assessment
2	**What is the relationship?** *What is the relationship between fear of falling incidence in African-American elders who live alone in community- based settings?*	Survey Correlation/passive Observation Ex post facto
3	**Why?** **cause/effect** *E.g. Does a program of fall-education reduce The incidence of falls in African-American elders Who live in community based settings*	Experimental design Quasiexperimental

Exercise on study design for registrars

For each of the following:

1. *What kind of study design is this? (there may be more than one answer, depending on how you look at it)*
2. *Are the objectives descriptive, analytic or diagnostic (or combinations)?*

1. For their MMed, a registrar takes all the patients being seen in the general surgery OPD over a 1-month period and interviews them on their satisfaction with their surgery. She reports the prevalence of satisfaction with the procedure, as well as factors associated with increased levels of satisfaction.

2. A Psychiatry registrar is interested in HIV infection among new admissions at Lentegeur. She uses an existing database to examine the prevalence of testing and (among those testing) HIV infection among patients admitted, and examines whether the prevalence of testing and/or HIV infection varies according to the nature of the admission.

3. A Medical registrar talks to a senior professor and decides to do his MMed analysing data from patient folders. All the patients were referred to GHS with Addison's disease over a 20 year period; he is reporting on their presentation, management and outcomes.

4. An Ophthalmology registrar is working with an existing dataset on outcomes of cataract surgery. Patients who underwent one of four different cataract procedures were followed for 2 years post-surgery, and the incidence of complications was compared between different types of surgery.

5. For her MMed, an Obs/Gynae registrar reviews the obstetric outcomes of women admitted with antepartum haemorrhage at GSH over the last 2 years. She is interested in the frequency of ICU admission in this group, and the factors associated with ICU admission.

6. For his MMed, an ENT registrar decides to study the validity of ENT registrars in diagnostic otitis media. She shows the same panel of 20 children (half of whom are known to have otitis media) to each of 50 registrars, and assesses the sensitivity and specificity of registrars' diagnoses.

Options in study design

- – Case report  Not acceptable for MMed
- – Case series  e.g. retrospective/XS
- – Case - Control  vs  comparison

- – Cohort Longitudinal
- - Lab study: tissue, blood, other body fluids, animals

Statistical Methods to Test Hypotheses

Type of experiment					
Scale of measurement	Two treatment groups consisting of different individuals	Three or more treatment groups consisting of different individuals	Before and after a single treatment in the same individuals	Multiple treatments in the same individuals	Association between two variables
Interval (and drawn From normally disturbed Populations)*	Unpaired t test (Chap. 4)†	Analysis of variance (Chap.3)	Paired t test (Chap.9)	Repeated-measures analysis of variance (Chap.9)	Linear regression And Pearson product-moment correlation; Bland-Altman Analysis (Chap.8)
Nominal	Chai-square analysis-of-contingency table (Chap.5)	Chai-square analysis-of-contingency table (Chap.5)	McNemar's test (Chap.9)	Cochrane Q†	Contingency coefficients†
Ordinal	Mann-Whitney Rank-sum test (Chap.10)	Kruskal-Wallis Statistic (Chap.10)	Wilcoxon signed-rank Test (Chap.10)	Friedman statistic (Chap.10)	Spearman rank Correlation (Chap.8)
Survival time	Log-rank test Or Gehan's test (Chap.11)				

***If the assumption of normally distributed populations is not met, rank observations and use methods of data measured on an ordinal scale.**
†Not included in this text.

THE UNIVERSITY of York

NHS CENTRE FOR REVIEWS AND DISSEMINATION

Undertaking Systematic Reviews of Research on Effectiveness

CRD's Guidance for those Carrying Out or Commissioning Reviews

CRD Report Number 4 (2nd Edition)

March 2001

Click on the links below to view each section of the report in pdf format. You will need a copy of Adobe Acrobat Reader®. Go to the Adobe web site to download a free copy of Acrobat Reader®.

To purchase a copy of this report (price £12.50), please contact the CRD Publications Office

---

CONTENTS

Click here for the CRD home page

Click here for CRD Publications

Systematic Review – Principles

A systematic review is a way of exploring the evidence on a particular issue. It involves a: -

1. Clearly formatted question to be addressed
2. "Exhaustive" search of the literature
3. Explicit pre-defined inclusion criteria for studies
4. Exploration of the quality of the papers included

Your preparatory work should therefore be spent on deciding upon these 4 things.

What exactly should I explore? Remember to keep the area small, and focussed.

Start to consider ways of searching the databases, find out how to locate and use the appropriate web based search engines, Medline, Cinahl, optonet etc.

Read up on issues such as reliability, validity, bias etc. You need this to look at the quality of the papers.

Reasons for doing a systematic review

There are several good reasons to conduct a systematic review:

- To gain a consensus of research evidence
- To find and summarise unmanageable amounts of data (this one you will **not** be doing in your dissertation.)
- To make sense of contradictory data. To be able to objectively explore differences in the effects of a treatment e.g. does it work or not? Or to examine which treatment is best?)
- To avoid bias in the interpretation of data (through prejudice, unconscious bias, incomplete knowledge or inappropriate weighing of the data).
- To improve the precision in estimating the effect of an intervention (meta-analysis is an example of this. Don't worry if you don't know what this means, its just a way of amalgamating the statistics of several studies or more).

Like any other type of research a systematic review can be done well or badly.
Good reviews are those which are transparent in their methods, you can see how they decided on good and bad papers, you can see how conclusions were reached etc.

A systematic review is very dependent on the data that it analyses. So you must be very careful in the process you use to collect and interpret the papers you use.

You must develop a method to screen out poor papers.

There are several very useful web sites on systematic reviews. They focus on reviews on a scale, which is well out of the scope of your dissertation, but they do provide lots of tips and ideas, references etc that will help you do it.

NHS centre for Reviews and dissemination
www.york.ac.uk/inst/crd/welcome.htm

From this site you can get access to The Cochrane Collaboration website and many other useful resources.

Please don't contact them to say your doing a review because the scale on which you are doing it is much too small for them. Most of their reviews involve research teams and are approx 2 years in duration. Do use their websites for information and guidance and for some nice examples of reviews that have been conducted. This could help a lot in the planning stages.

Volume 1, number 5

What is a systematic review?

Sponsored by an educational grant from Aventis Pharma

Huw T O Davies
PhD HonMFPHM
Lecturer in Health Care Management, University of St Andrews

and

Iain K Crombie
PhD HonMFPHM
Reader in Epidemiology, University of Dundee

- **Systematic reviews are superseding narrative reviews** as a way of summarising research evidence.
- Systematic reviews attempt to bring the **same level of rigour to reviewing research evidence** as should be used in **producing that research evidence** in the first place.
- High-quality systematic reviews take great care to **find all relevant studies** published and unpublished, **assess each study, synthesise the findings from individual studies in an unbiased way** and **present a balanced and impartial summary** of the findings with due consideration of any flaws in the evidence.
- **Many high-quality reviews are available** both in journals and from electronic sources such as the Cochrane Library.
- Not all published systematic reviews have been produced with meticulous care – therefore the **findings may sometimes mislead**. Interrogating published reports by asking **a series of questions can uncover deficiencies**.

Abbreviated prescribing information is on page 6

What are systematic reviews?

The need for reviews

The explosion in biomedical publishing in the latter half of the 20th century (perhaps 20,000 journals and upwards of 2 million articles a year) makes keeping up with primary research an impossible feat.

Moreover, clinicians, therapists and healthcare managers have wide-ranging information needs – that is, they need good information on the effectiveness of a large number of therapeutic interventions; not just one or two.

In even a single area it is not unusual for the number of published trials to run into dozens or even hundreds. Further, many of these studies will give unclear, confusing or downright contradictory results. Looked at individually, each trial may offer little insight into effectiveness; the hope is that, when taken together, a clearer (and more consistent) picture will emerge.

Failings in traditional reviews

Reviews have always been a part of the medical literature. Respected peer leaders, experts in their field, have sought to collate existing knowledge and publicise these summaries. Frequently such reviews have been about assessing the effectiveness (or otherwise) of therapeutic interventions.

Unfortunately, such attempts at synthesis have not always been as rigorous as might have been hoped. One obvious problem is that, traditionally, reviewers rarely began with an open mind as to the likely recommendations. Indeed, those involved in developing a review may well have started a review (or have been commissioned to write one) precisely because of their accumulated experience and professional opinions.

However, if strong prior beliefs are held then a dispassionate review of evidence will be difficult to achieve. At worst, a reviewer may simply build a case in support of their personal beliefs, selectively citing appropriate studies along the way. Even if the reviewer does begin with an open mind, traditional narrative reviews are rarely explicit about how studies are selected, assessed and integrated. Thus the reader is generally unable to assess the likelihood of prior beliefs or other biases clouding the review process.

For all this, such narrative reviews were and are widespread and influential.

The lack of rigour in the creation of reviews went largely unremarked until the late 1980s when several commentators exposed the inadequacies of the process and the consequent bias in recommendations.[1,2] Not least of the problems was that small but important effects were being missed, different reviewers were reaching different conclusions from the same research base and the findings reported often had more to do with the specialty of the reviewer than with the underlying evidence.[3]

The inadequacy of traditional reviews and the need for a rigorous systematic approach were emphasised in 1992 with the publication of two landmark papers.[4,5] In these Elliot Antman, Joseph Lau and colleagues reported two devastating findings:

- **First, that if original studies of the effects of clot busters after heart attacks had been systematically reviewed the benefits of therapy would have been apparent as early as the mid-1970s.**
- Second, Antman and Lau showed that **text books and narrative reviews were woefully inadequate in summarising the current state of knowledge**. These reviews either omitted mention of effective therapies or suggested that the treatments should be used only as part of an ongoing investigation – when in fact the evidence (if it had been collated) was near incontrovertible.

These papers showed that there was much knowledge to be gained from collating existing research, but that traditional approaches had largely failed to extract this knowledge. What was needed was the same rigour in secondary research (research where the objects of study

A review of only leading journals is likely to give an over-optimistic view of therapy effectiveness

are other research studies) as is expected from primary research (original trials).

When systematic reviews are needed

Systematic reviews are needed whenever there is a substantive therapeutic question, several primary studies – perhaps with disparate findings – and substantial uncertainty.

For example, the growing literature comparing the antithrombotic action of low molecular-weight heparins with unfractionated heparin[6,7] suggests that there may be some benefits in using the newer drugs. However, a recent narrative review[8] enumerated a number of unanswered therapeutic questions and sounded an important caveat: not all low molecular-weight heparins are the same. Here, then, is an area where a systematic review may help clarify matters; preparing one, however, is not a trivial exercise.

The process of systematic review

The need for rigour in the production of systematic reviews has led to the development of a formal process for their conduct. Understanding the approach taken and the attempts to minimise bias can help in the appraisal of published reviews.

Briefly, developing a systematic review requires the following steps:

1. Defining an appropriate therapeutic question. This requires a clear statement of the intervention of interest, relevant patient groups (and sometimes the settings where the intervention is administered), as well as appropriate outcomes. These details are used to select studies for inclusion in the review.

2. Searching the literature. The published and unpublished literature are carefully searched for all reports of controlled trials of this intervention (on the right patients, reporting the right outcomes and so on). For an ***unbiased*** assessment, this search must cover all the literature (not just Medline, where typically less than half of all trials will be found), including non-English sources. Further, studies reported only at conferences, in company reports or unreported and buried in filing cabinets must also be sought.

The concern is over ***publication bias***[9,10] – the notion that studies which report a positive effect of therapies are more likely to be reported in good English language journals than studies that report no effect. Thus a review of only leading journals is likely to give an overoptimistic view of therapy effectiveness.

3. Assessing the studies. Once all possible study reports have been identified, each study needs to be assessed for **eligibility** for inclusion, **study quality** and **reported findings**. Ideally, such assessment should involve two independent reviewers.

4. Combining the results. The findings from the individual studies must then be aggregated to produce a 'bottom line' on the clinical effectiveness of the intervention. Sometimes this aggregation is qualitative, but more usually it is a quantitative assessment using a technique known as meta-analysis (see ***What is meta-analysis?*** in this series).

5. Placing the findings in context. The findings from this aggregation of an unbiased selection of studies then need to be discussed to put them in context. This will address such issues as the quality and heterogeneity of the included studies, the likely impact of bias and chance and the applicability of the findings. Thus judgement and balance are not obviated by the rigour of systematic reviews – they are just reduced in impact and made more explicit.

Useful websites for systematic reviews

Systematic Reviews Training Unit
http://www.ich.ucl.ac.uk/srtu

Cochrane Collaboration
http://hiru.mcmaster.ca/cochrane/default.htm

NHS Centre for Reviews & Dissemination
http://www.york.ac.uk/inst/crd/welcome.htm

Centre for Evidence-Based Medicine at Oxford
http://cebm.jr2.ox.ac.uk/

Bandolier
http://www.jr2.ox.ac.uk/bandolier/index.html

A word of caution, however. Performing a rigorous systematic review is far from easy. It requires meticulous and laborious searching and considerable attention to methodological detail before it truly deserves the badge 'systematic'. Clear guidance on the process of developing systematic reviews is available[11,12] as are courses run at Oxford and other centres of excellence.

Finding existing reviews

High-quality systematic reviews are published in many of the leading journals. In addition, electronic publication by the Cochrane Collaboration, the NHS Centre for Reviews and Dissemination and others offers speedy access to regularly updated summaries (see Box, page 3).

Drawbacks of systematic reviews

Systematic reviews appear at the top of the 'hierarchy of evidence' (see Box, below). This reflects the fact that, when well conducted, they should give us the best possible estimate of any true effect. As noted previously, such confidence can sometimes be unwarranted, however, and caution must be exercised before accepting the veracity of any systematic review. A number of problems may arise:

- First, like any piece of research, **a systematic review may be done badly**. Attention to the questions listed in the section 'Appraising a systematic review' can help separate the rigorous research from the quick and slapdash.
- **Inappropriate aggregation of studies** that differ in terms of intervention used or patients included can lead to the **drowning of important effects.** For example, the effects seen in some subgroups may be concealed by a lack of effect (or even reverse effects) in other subgroups.
- The **findings from systematic reviews are not always in harmony with** the findings from **large-scale high-quality single trials**.[13,14] Thus findings from systematic reviews need to be weighed against perhaps conflicting evidence from other sources. Ideally, an updated review would deal with such anomalies.

Appraising a systematic review

Not all systematic reviews are rigorous and unbiased. The reader will want to interrogate any review that purports to be systematic to assess its limitations. The following questions provide a framework. Further guidance on appraising a systematic review can be found in several useful publications.[15–17]

- **Is the topic well defined** in terms of the intervention under scrutiny, the patients receiving the intervention (plus the settings in which it was received) and the outcome that was assessed?
- **Was the search for papers thorough?** Was the search strategy described? Was manual searching used as well as electronic databases? Were non-English sources searched? Was the 'grey literature' covered – for example, non-refereed journals, conference proceedings or unpublished company reports? What conclusions were

Hierarchies of evidence

I–1 Systematic review of several double-blind randomised control trials.

I–2 One or more large double-blind randomised control trials.

II–1 One or more well-conducted cohort studies.

II–2 One or more well-conducted case-control studies.

II–3 A dramatic uncontrolled experiment.

III Expert committee sitting in review; peer leader opinion.

IV Personal experience.

drawn about the possible impact of publication bias?

- **Were the criteria for inclusion of studies clearly described and fairly applied?** For example, were blinded or independent reviewers used?
- **Was study quality assessed by blinded or independent reviewers?** Were the findings related to study quality?
- **Was missing information sought from the original study investigators?** Was the impact of missing information assessed for its possible impact on the findings?
- **Do the included studies seem to indicate similar effects?** If not, was the heterogeneity of effect investigated, assessed and discussed?
- **Were the overall findings assessed for their robustness** in terms of the selective inclusion or exclusion of doubtful studies and the possibility of publication bias?
- **Was the play of chance assessed?** In particular, was the range of likely effect sizes presented and were null findings interpreted carefully? That is, a review that finds no evidence of effect may simply be an expression of our lack of knowledge rather than an assertion that the intervention is worthless.
- **Are the recommendations based firmly on the quality of the evidence presented?**

In their enthusiasm, reviewers can sometimes go beyond the evidence in drawing conclusions and making their recommendations.

Conclusion

All studies have flaws. It is not the mere presence of flaws that vitiates the findings. Even flawed studies may carry important information. The reader must exercise judgement in assessing whether individual flaws undermine the findings to such an extent that the conclusions are no longer adequately supported.

References

1. Mulrow CD. The medical review article: state of the science. *Ann Intern Med* 1987; **106:** 485–488.
2. Teagarden JR. Meta-analysis: whither narrative review? *Pharmacotherapy* 1989; **9:** 274–284.
3. Spector TD, Thompson SG. The potential and limitations of meta-analysis. *J Epidemiol Community Health* 1991; **45:** 89–92.
4. Antman EM, Lau J, Kupelnick B, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. *JAMA* 1992; **268:** 240–248.
5. Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992; **327:** 248–254.
6. Cohen M, Demers C, Gurfinkel EP *et al.* A comparison of low-molecular-weight heparin with unfractionated heparin for unstable coronary artery disease. *N Engl J Med* 1997; **337:** 447–452.
7. Levine M, Gent M, Hirsh J *et al.* A comparison of low-molecular-weight heparin administered primarily at home with unfractionated heparin administered in the hospital for proximal deep-vein thrombosis. *N Engl J Med* 1996; **334:** 677–681.
8. Armstrong PW. Heparin in acute coronary disease – requiem for a heavyweight? *N Engl J Med* 1997; **337:** 492–494.
9. Egger M, Zellweger-Zähner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. *Lancet* 1997; **350:** 326–329.
10. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991; **337:** 867–872.
11. Chalmers I, Altman DG (eds). *Systematic Reviews*. London: BMJ Publishing Group, 1995.
12. Cook DJ, Sackett DL, Spitzer WO. Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam consultation on meta-analysis. *J Clin Epidemiol* 1995; **48:** 167–171.
13. Le Lorier J, Gregoire G, Benhaddad A, Lapierre J, Derderian F. Discrepancies between meta-analyses and subsequent large randomized controlled trials. *N Engl J Med* 1997; **337:** 536–542.
14. Egger M, Davey-Smith G. Misleading meta-analysis. Lessons from 'an effective, safe, simple' intervention that wasn't. *BMJ* 1995; **310:** 752–754.
15. Crombie IK. *The Pocket Guide to Critical Appraisal.* London: BMJ Publishing, 1996.
16. Milne R, Chambers L. Assessing the scientific quality of review articles. *J Epidemiol Community Health* 1993; **47:** 169–170.
17. Oxman AD, Cook DC, Guyatt GH. Users' guides to the medical literature: VI. How to use an overview. *JAMA* 1994; **272:** 1367–1371.

Volume 4, number 1

What is clinical audit?

Sponsored by an educational grant from Aventis Pharma

Tracey Jones MSc
Clinical Audit and Effectiveness Manager, North Bristol NHS Trust

Simon Cawthorn
BSc MBBS MS FRCS
Consultant Surgical Oncologist, North Bristol NHS Trust

- Clinical audit is a quality improvement process that aims to improve patient care and outcomes by carrying out a systematic review and implementing change. Aspects of patient care – including structure, processes and outcomes – are selected and evaluated against explicit criteria and, where necessary, changes are implemented at an individual, team or service level. Further monitoring can then be used to confirm the improvements in healthcare delivery. This definition is endorsed by the National Institute for Clinical Excellence (NICE).
- Clinical audit provides the framework to improve the quality of patient care in a collaborative and systematic way, as outlined in current NHS policy statements.
- The report of the public inquiry into children's heart surgery at the Bristol Royal Infirmary 1984–1995 (2001) highlights the importance of clinical audit.
- Clinical governance presents a new challenge – to take audit 'at its best' and incorporate it within organisation-wide approaches to quality (see *What is clinical governance?*).
- Topics for audit projects should reflect national and/or local targets; for example, in cancer services, coronary care or mental health. Projects may also need to focus on the implementation of National Service Frameworks (NSFs), Health Improvement and Modernisation Plans (HIMPs) or NICE guidelines and appraisals.
- Clinical audit has a mixed history in the NHS. For it to become an important component in the management of health services, a change needs to take place in the standing of audit programmes. Audit can no longer be seen as a fringe activity for enthusiasts within clinical governance. Instead, the NHS needs to make a commitment to support audit as a mainstream activity.
- Clinical audit, when it is conducted well, provides a way in which the quality of care can be reviewed objectively, within an approach which is supportive and developmental.

Prescribing information is on page 8

What is clinical audit?

Clinical audit should be an integral part of clinical practice. All clinicians want to provide the best possible care for patients, and clinical audit is one tool that can help this to happen in a systematic way. It can be a powerful tool for positive change, resulting in improved practice and outcomes for patients, and it has become integral to NHS policy.

- As a first step, clinical audit was integrated into clinical governance systems.[1,2] (See *What is clinical governance?*)[3]
- Full participation in clinical audit by all hospital doctors was subsequently made an explicit component of clinical governance.[2,4]
- The NHS Plan[5] took these policies further, with proposals for mandatory participation by all doctors in clinical audit, and support for the involvement of other staff, including nurses, midwives and therapists. Meanwhile, *Improving health in Wales*[6] introduced annual appraisals to address the results of audit.
- Clinical audit in Scotland is a multiprofessional activity which is supported locally, but also nationally, by the Clinical Resource and Audit Group (CRAG) via effectiveness programmes and national audits. Standards for audit are increasingly being drawn from recommendations in the Scottish Intercollegiate Guidelines Network (SIGN) guidelines, and from topics selected by the Clinical Standards Board for Scotland (CSBS) for quality assurance.[7] Current plans are that the CSBS, along with Scotland's two other central clinical effectiveness organisations (the Health Technology Board for Scotland and the Scottish Health Advisory Service), will be integrated into a single new health board, NHS Quality Improvement Scotland.

The report of the public inquiry into children's heart surgery at the Bristol Royal Infirmary 1984–1995 (2001) highlights the importance of clinical audit.[8] The General Medical Council now advises all doctors that they 'must take part in regular and systematic medical and clinical audit, recording data honestly. Where necessary, you must respond to the results of audit to improve your practice, for example by undertaking further training'.[9] The Nursing and Midwifery Council (formerly the UKCC) stated that in clinical governance, assisting the co-ordination of quality improvement initiatives is 'the business of every registered practitioner'.[10]

The history of clinical audit

Clinical audit has a mixed history in the NHS. Many projects have been run into the ground without having demonstrated much of a contribution to the quality of services – but there have also been significant successes. Many local projects have provided a systematic structure through which clinical teams can deliver real improvements in patient care. In some cases, national projects have been able to play an important role in service-wide changes, improving access and quality of care throughout the country (the national audit of stroke care is the most well-known example).[11]

In 1989, the White Paper *Working for patients* defined medical audit as 'the systematic critical analysis of quality of medical care including the procedures used for diagnosis and treatment, the use of resources and the resulting outcome and quality of life for the patient'.[12]

Medical audit later evolved into clinical audit and a revised definition was announced by the NHS Executive: 'Clinical audit is the systematic critical analysis of the quality of healthcare, including the procedures used for diagnosis, treatment and care, the use of resources and the resulting outcome and quality of life for the patient'.[13]

NICE recently published *Principles for best practice in clinical audit*,[11] which defines clinical audit as 'a quality improvement process that seeks to improve patient care and outcomes through systematic review of care against explicit criteria and the implementation of change. Aspects of the structure, processes, and outcomes of care are

selected and systematically evaluated against explicit criteria. Where indicated, changes are implemented at an individual, team, or service level and further monitoring is used to confirm improvement in healthcare delivery'.

Box 1. Making clinical audit work[15,16]

Audit cycle of the marking of breast wide local excision specimens
An audit was set up by Dr Lucy James at the Frenchay Breast Care Centre in Bristol to identify whether all breast wide local excision (WLE) specimens were being marked correctly prior to histopathological examination. The audit resulted in the introduction of a simple sticker as an adjunct to the standard pathology form, which improved standards for the marking of breast WLE specimens by 64.4%.

Marking of this type of specimen not only aids the pathologist in orientating the tumour within the surgical specimen, but also allows the surgeon to identify exactly which border requires further surgery if the histological examination reveals the tumour to be too near the excision margins. The greatest impact may be for the patient, however, as the cosmetic result will be more favourable if subsequent surgery is as conservative as possible.

Policy objective measured
- Reduce death/complications from cancers and their treatment

Indicator title
- NB1 4 node status
- Percentage of cases in which histological node status was ascertained
- Percentage of cases in which histological node status was ascertained by clearance or sampling of at least four nodes

Performance assessment framework (PAF) domain
- Effective delivery of appropriate healthcare (ED)

References
- COG recommendation 100%
- BASO standard 100%

Touch preparation cytology from breast core biopsies for one-stop diagnosis
Ultrasound-guided biopsy is rapidly replacing cytological diagnosis in the triple assessment of solid breast lumps. Touch preparations of 112 consecutive patients undergoing ultrasound-guided core biopsies were taken for instant cytological assessment. The cytological grades of these touch preparations were compared retrospectively with the core histological diagnoses. Results gave a sensitivity of 91%, a specificity of 95%, a positive predictive value of 99% and a negative predictive value of 72%, suggesting that touch preparation from ultrasound-guided core biopsy of solid breast lumps provides a reliable, accurate diagnosis in a one-stop clinic.

Policy objective measured
- Reduce death/complications from cancers and their treatment

Indicator title
- NB1 6 triple assessment
- Percentage of patients with breast cancer receiving triple assessment on first visit

PAF domain
- ED
- Fair access (FA)
- Efficiency (EF)
- Patient/carer experience (PEx)

References
- COG (should be available on first visit)
- BASO (less than 10% of patients should have to attend the hospital on more than two occasions for diagnostic purposes)

Types of audit

Standards-based audit – A cycle which involves defining standards, collecting data to measure current practice against those standards, and implementing any changes deemed necessary.

Adverse occurrence screening and critical incident monitoring – This is often used to peer-review cases which have caused concern or from which there was an unexpected outcome. The multidisciplinary team discusses individual, anonymous cases to reflect upon the way the team functioned and to learn for the future. In the primary care setting, this is described as a 'significant event audit'.

Peer review – 'An assessment of the quality of care provided by a clinical team with a view to improving clinical care'.[14] Individual cases are discussed by peers to determine, with the benefit of hindsight, whether the best care was given. This is similar to the method described above, but might include 'interesting' or 'unusual' cases rather than problematic ones. Unfortunately, recommendations made from these reviews are often not pursued as there is no systematic method to follow.

Patient surveys and focus groups – These are methods used to obtain users' views about the quality of care they have received. Surveys carried out for their own sake are often meaningless, but when they are undertaken to collect data they can be extremely productive.

Selecting an audit project

The clinical team has an important role in prioritising clinical topics, and the following questions may be a useful discussion guide.[11]
- Is the topic related to high cost, volume or risk to staff or users?
- Is there any evidence of a serious quality problem; for example, patient complaints or high complication rates?
- Is good evidence available to inform standards; for example, systematic reviews or national clinical guidelines?
- Is the problem amenable to change?
- Is sustainable improvement possible?
- Is there any potential for involvement in a national audit project?
- Is the topic pertinent to national policy initiatives?
- Is the topic a priority for the organisation?

What is clinical audit?

Box 2. The role of audit in the implementation of NICE guidance

In December 2001, the government issued directions making it mandatory for health authorities to act on NICE recommendations. Clinical audit programmes should now record the proportion of treatments adhering to NICE guidance.

Example 1: NICE Technology Appraisal Guidance No. 33: 'Guidance on the use of irinotecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer'[17]

In March 2002, NICE issued guidance to the NHS on the use of selected therapies for advanced colorectal cancer. The guidance stated the following.

- On the balance of clinical and cost-effectiveness, neither irinotecan nor oxaliplatin in combination with 5-FU/FA are recommended for routine firstline therapy for advanced colorectal cancer.
- Oxaliplatin in combination with 5-FU/FA should be considered as firstline therapy in advanced colorectal cancer in patients with metastases that are confined solely to the liver and may become resectable (down-staged) following treatment.
- Irinotecan monotherapy is recommended in patients who have failed an established 5-FU-containing treatment regimen.
- On the balance of evidence relating to clinical and cost-effectiveness, raltitrexed is not recommended for the treatment of advanced colorectal cancer.

Example 2: NICE Technology Appraisal Guidance No. 30: 'Guidance on the use of taxanes for the treatment of breast cancer'[19]

In September 2001, NICE reviewed its guidance for taxanes in breast cancer and reinstated the original guidance that was issued in June 2000. The guidance stated the following.

- The use of docetaxel in combination with an anthracycline in firstline treatment of advanced breast cancer is not currently recommended. As paclitaxel is not licensed for firstline use with anthracycline, its use has not been considered in this indication.
- Docetaxel and paclitaxel are recommended as an option for the treatment of advanced breast cancer where initial cytotoxic chemotherapy (including an anthracycline) has failed or is inappropriate.
- The taxanes are not currently licensed in the UK for adjuvant treatment of early breast cancer and their use in this indication should, therefore, be limited to randomised clinical trials.

In many NHS organisations, a committee or clinical effectiveness/ governance team decides which clinical audit projects should be undertaken each year.[11] Their decisions are usually based on local health priorities which reflect national targets; for example, in cancer services, coronary care or mental health (see Box 1, page 3).[11] Projects may also need to focus on the implementation of NSFs, HIMPs or NICE guidelines and appraisals.[11] For example, in March 2002 NICE issued guidance on the use of selected therapies for advanced colorectal cancer (see Box 2).[17]

The *Breast Cancer Service Guidance*, published by NICE, outlines evidence-based recommendations and appropriate tools and techniques for measurement.[18] For example, in the management of advanced, recurrent and metastatic disease it proposes that patients who do not receive taxanes should have the reason recorded.[18,19] This reflects the mandatory nature of NICE guidance.

Clinical audit – the process

Clinical audit can be described as a cycle or a spiral – see Figure 1 (opposite).[11] Within the cycle there are stages that follow the systematic process of: establishing best practice; measuring care against criteria; taking action to improve care; and monitoring to sustain improvement.[11] As the process continues, each stage aspires to a higher level of quality.[11]

Figure 2 (page 6) outlines the five stages of clinical audit, which involves the use of specific methods, but also requires the creation of a supportive environment.[11]

Stage 1: Preparing for audit

National audit projects reviewed by NICE suggest that two broad areas of preparation must be addressed:[11]

- Project management, including topic selection, planning and resources, and communication
- Project methodology, including design, data issues, ease of implementation, stakeholder involvement, and the provision of support for local improvement.

In practical terms, preparing for audit can be broken down into five elements:[11]

- Involving users in the process ('users' include patients, carers and the groups and organisations that represent their interests)
- Selecting a topic
- Defining the purpose of the audit
- Providing the necessary structures
- Identifying the skills and people needed to carry out the audit, and training staff and encouraging them to participate.

Stage 2: Selecting criteria/standards

In clinical audit, criteria or standards are used to assess the quality of care provided by an individual, a team or an organisation.[11] These criteria are explicit statements that define what is being measured and represent elements of care that can be measured objectively.[11]

Recent government publications indicate that health professionals will be expected to develop criteria and standards that measure a wide range of aspects of quality, such as access to care and patient satisfaction.[5,11]

Criteria can be classified into those concerned with:[11]

- Structure (what you need)
- Process (what you do)
- Outcome of care (what you expect).

Stage 3: Measuring performance

To ensure that the data collected are precise, and that only essential data are collected, certain details of what is to be audited must be established from the outset.[11] These are:[11]

- The user group to be included, with any exceptions noted
- The healthcare professionals involved in the users' care
- The period over which the criteria apply.

Sampling is sometimes a contentious issue in audit. It is necessary first to define the population to which the audit applies; for example, all women presenting to the breast clinic during a specific year. However, it might be impractical to collect data on every single woman in the population, so a representative sample can be used instead. The necessary sample size can be calculated scientifically, as long as it is possible to estimate the proportion of patients who are likely to meet the criteria, and the level of confidence that is required in the results. Pragmatic sample sizes are more commonly used because of limitations in timescale and resources. A time frame is often used to define the sample: for example, all women referred to the breast clinic in a one-month period. Alternatively, a consecutive sample of patients might be used; for example, the last 100 referrals. It is important that all those likely to be affected by the audit results agree on the sample sizes and agree that they will act on the results.

Data are collected in order to measure current practice against agreed standards. Some data may already be available in computerised clinical information systems. In other cases, appropriate data may have been collected routinely by other methods. Where data are not routinely collected, or are held only in paper records, it may be necessary to devise a data collection form on which to record information (see Figure 3, page 7).[11] The data collected should relate only to the objectives of the audit – do not be tempted to collect additional, 'interesting' information. Respect patient and staff confidentiality; identifiable information should not be used. Ethical consideration should also be given to the design of the project, and any potentially sensitive topics should be discussed with the local Research Ethics Committee, particularly where patients' views are to be sought.

If the data collection strategy includes asking users/carers for their views, care must be taken in developing a questionnaire. It is a good idea to pilot the form on a number of users to ensure that it works. Simple statistical analysis is usually all that is required of audit data. Were the standards met? If not, why not?

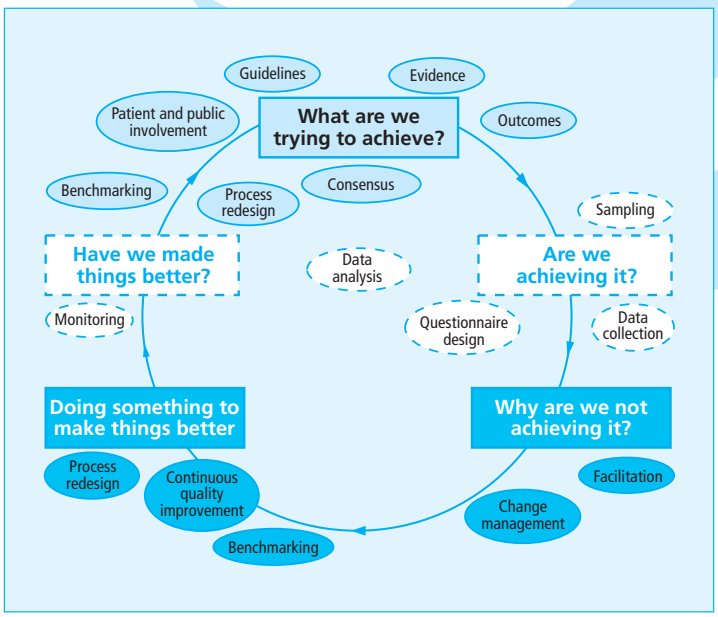


Figure 1. The clinical audit cycle[11]

Stage 4: Making improvements

Once the results of the audit have been published and discussed, an agreement must be reached about the recommendations for change. Using an action plan to record

these recommendations is good practice; this should include who has agreed to do what, and by when. Each action point needs to be well defined, with an individual named as responsible for it, and an agreed timescale for its implementation.

People have lots of good reasons to change ... and not to change. There are many strategies for altering behaviour to enable change. Kurt Lewin's famous model of 'unfreeze' (obtaining consensus that a change is required), 'move' (making the change); and 'refreeze' (ensuring the change is embedded in practice) is a useful one.[20] The best way to get everyone to agree that the change will be beneficial is to ensure that they participate in the whole process.

Aside from those individuals who will never agree to the need for change – regardless of the evidence – there are potential barriers to change in terms of resources, politics or environment. Change needs to be implemented in a systematic way, ensuring that communication and dissemination are sustained throughout the process.

Stage 5: Sustaining improvements

After an agreed period, the audit should be repeated. The same strategies for identifying the sample, methods and data analysis should be used to ensure comparability with the original audit. The re-audit should demonstrate that the changes have been implemented and that improvements have been made. Further changes may then be required, leading to additional re-audits.

Evidence suggests that multifaceted change strategies achieve the optimum maintenance of improvements. Feedback of results to clinicians (anonymously) can highlight achievements, while training and education help to ensure that new practice is embedded as routine. Dissemination of guidelines or protocols is useful if accompanied by training and awareness campaigns, and decision support or reminders may also help to introduce new practices effectively.

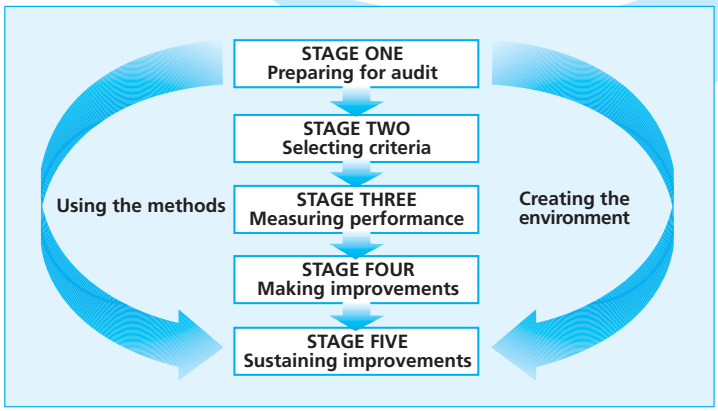


Figure 2. The stages of clinical audit[11]

Conclusion

Increasing multiprofessional participation is the key to a successful audit. Successful audit means good-quality studies that are based on agreed, evidence-based standards of care, that have agreed outcomes and that achieve sustained improvements in care for patients.

Further reading

Clare Morrell, Gill Harvey. Royal College of Nursing. *The clinical audit handbook. Improving the quality of health care.* London: Harcourt Brace & Co Ltd, 1999.

References

1. Department of Health. *The new NHS: Modern. Dependable.* London: The Stationery Office, 1997.
2. Welsh Office. *Quality care and clinical excellence.* Cardiff: Welsh Office, 1998.
3. Starey N. *What is clinical governance?* London: HMC, 2001.
4. Department of Health. *A first class service. Quality in the new NHS.* London: DoH, 1998.
5. Department of Health. *The NHS Plan: a plan for investment – a plan for reform.* London: The Stationery Office, 2000.
6. Minister for Health and Social Services. *Improving health in Wales – a plan for the NHS and its partners.* Cardiff: National Assembly for Wales, 2001.
7. Scottish Executive. *Focus on quality.* March 2002.
8. Department of Health. *Learning from Bristol: The Department of Health's Response to the Report of the Public Inquiry into the children's heart surgery at the Bristol Royal Infirmary 1984–1995.* London: The Stationery Office, 2002.
9. General Medical Council. *Good medical practice.* London: General Medical Council, 2001.
10. UK Central Council for Nursing, Midwifery and Health Visiting. *Professional self-regulation and clinical governance.* London: United Kingdom Central Council for Nursing, Midwifery and Health Visiting, 2001.
11. National Institute for Clinical Excellence. *Principles for best practice in clinical audit.* Oxford: Radcliffe Medical Press, 2002.
12. Department of Health. *Working for patients.* London: The Stationery Office, 1989.
13. NHS Executive. *Promoting clinical effectiveness. A framework for action in and through the NHS.* London: NHS Executive, 1996.
14. Swage T. *Clinical governance in health care practice.* Oxford: Butterworth-Heinemann, 2000.
15. Data on file. Bristol NHS Trust.
16. NHS Executive. *Manual of cancer services standards.* London: NHS Executive, 2000.
17. NICE. *Guidance on the use of irinotecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer: National Institute for Clinical Excellence Technology Appraisal Guidance No 33.* London: NICE, 2002.
18. NICE. *Breast Cancer Service Guidance.* 2002. (www.nice.org.uk/cat.asp?c=36017) Last accessed 22 November 2002.
19. NICE. *Guidance on the use of taxanes for the treatment of breast cancer: National Institute for Clinical Excellence Technology Appraisal Guidance No 30.* London: NICE, 2001.
20. Lewin K. Group decision and social change. In: Swanson GE, Newcomb TM, Hartley EL (eds). *Readings in social psychology.* New York: Holt, Rinehart & Winston, 1952.

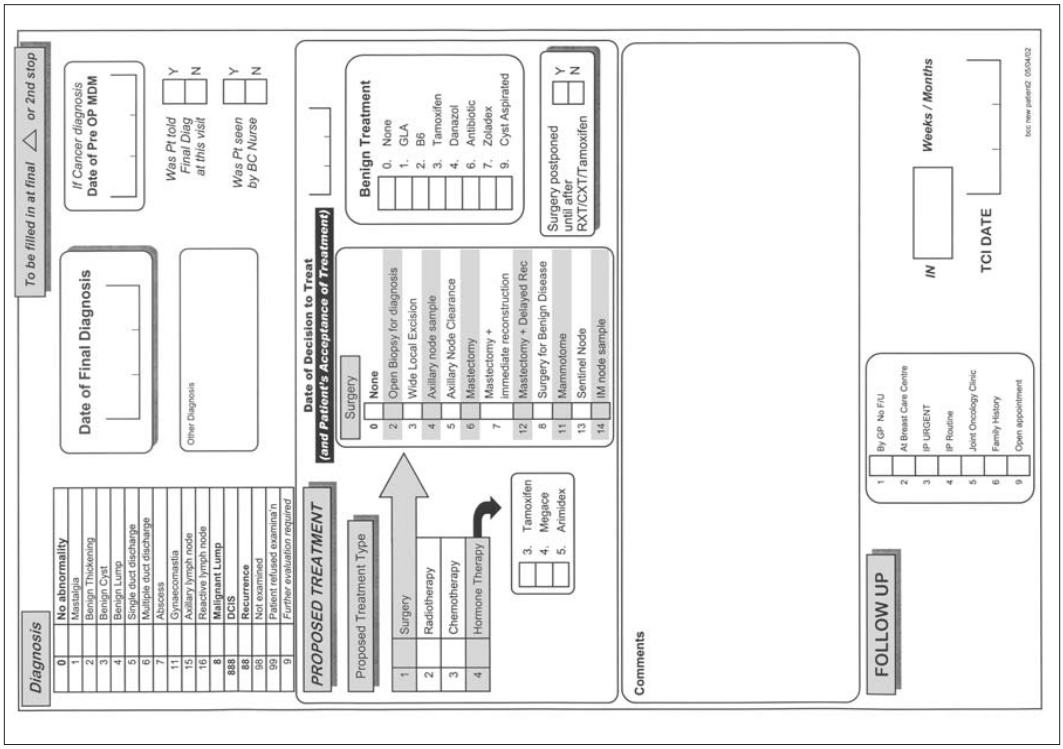


Figure 3. Example of a data collection form[15]

Volume 4, number 1

What is clinical audit?

Campto® (irinotecan hydrochloride trihydrate) Prescribing Information

Presentations: Vials of concentrate for infusion containing either 40 mg or 100 mg irinotecan hydrochloride trihydrate. **Indications:** Treatment of adult patients with advanced colorectal cancer. **Dosage & Administration:** Solution must be prepared aseptically. Campto should be administered as an intravenous infusion over 30 to 90 minutes. In first line: combination therapy of 180 mg/m^2 every 2 weeks followed by folinic acid and 5-fluorouracil; in second line: monotherapy 350 mg/m^2 every 3 weeks. Prophylactic anti-emetics are recommended. ***Dosage Adjustments:*** Subsequent cycles should follow appropriate recovery of all adverse events to grade 0 or 1 NCI-CTC and resolution of diarrhoea. Dose reduction of 15-20% recommended if patients experience grade 4 neutropenia, febrile neutropenia, thrombocytopenia, grade 4 leucopenia or grade 3-4 non-haematological toxicity ***Impaired hepatic function:*** Monitor liver function regularly. Blood bilirubin levels (up to 3 times ULN) in patients with performance status £2, should determine the starting dose of Campto. Bilirubin up to 1.5 times ULN the recommended dosage is 350mg/m^2. Bilirubin 1.5 – 3 times ULN the recommended dosage is 250mg/m^2. patients with bilirubin beyond 3 times ULN should not be treated with Campto. ***Impaired renal function:*** Not recommended. ***Elderly:*** Care due to the greater frequency of decreased biological functions. **Contraindications:** Chronic inflammatory bowel disease and/or bowel obstruction; severe hypersensitivity reactions to Campto; pregnancy; breastfeeding; severe bone marrow failure; WHO performance status > 2. **Warnings and Precautions:** Use in units specialised in the administration of cytotoxic chemotherapy and under the supervision of an oncologist. Patients needing closer follow-up or particular risk of neutropenia weekly dosing schedule (125 mg/m^2/week for 4 weeks, then 2 weeks rest) may be considered. Patients should be aware of the risks of acute cholinergic syndrome and neutropenia, and management of delayed diarrhoea (occuring > 24 hours after the infusion). Loperamide should not be given prophylactically. Weekly monitoring of full blood counts recommended. Patients should not drive if dizziness or visual disturbances occur. Contraceptive measures must be taken during and for 3 months after therapy. **Interactions:** Care in patients receiving neuromuscular blocking agents. **Adverse reactions:** Delayed diarrhoea (requires immediate treatment with loperamide). Uncommonly, pseudomembranous colitis. Neutropenia, fever, anaemia, thrombocytopenia, nausea and vomiting, acute cholinergic syndrome. Infrequently intestinal obstruction, ileus or gastrointestinal haemorrhage, intestinal perforation, and increases of amylase and/or lipase. Transient increases in transaminases, alkaline phosphatase, bilirubin or creatinine. Dyspnoea, muscular contractions, cramps, paraesthesia, asthenia, reversible alopecia, dehydration, constipation; infrequently dehydration- related renal insufficiency, hypotension or circulatory failure. Mild effects include anorexia, cutaneous reactions, abdominal pain, and mucositis. Uncommonly, allergy and infusion site reactions, transient speech disorders associated with Campto infusions. **Pharmaceutical Precautions:** Do not mix with any other medications. Complete infusion within 12 hours of reconstitution, if stored at room temperature (22 ± 4∞C) or 24 hours, if stored at 2 - 8 ∞C. Comply with prevailing cytotoxic handling guidelines when preparing or handling Campto. **Legal category:** POM **PL Number:** 40 mg: 0012/0302, 100 mg: 0012/0303. **Basic NHS Price:** Campto 40 mg; £53.00; Campto 100 mg; £130.00. Further information is available on request from Aventis Pharma Ltd, 50 Kings Hill Avenue, West Malling, Kent. ME19 4AH.
Last revision of text: Jan 2002.

Taxotere® (docetaxel) Prescribing Information

Presentation: Vials of concentrate for infusion containing 20mg docetaxel or 80mg docetaxel with accompanying vials of solvent. **Indications:** Locally advanced or metastatic breast cancer in combination with doxorubicin for patients who have not received prior cytotoxic therapy for this condition. Locally advanced or metastatic breast cancer after failure of cytotoxic therapy, which should have included an anthracycline or alkylating agent. Locally advanced or metastatic non-small cell lung cancer (NSCLC) after failure of prior chemotherapy. **Dosage and Administration:** Taxotere is administered as a one-hour iv infusion every three weeks. The recommended dosage in breast cancer is 100 mg/m^2, or 75 mg/m^2 in combination with doxorubicin (50 mg/m^2). The recommended dosage in NSCLC is 75 mg/m^2. Premedication with an oral corticosteroid is recommended for 3 days, starting one day prior to docetaxel administration. ***Elderly:*** No special instructions. ***Children:*** Safety and efficacy not established. ***Hepatic impairment:*** Reduce dosage; discontinue in severe cases.
Contraindications: Hypersensitivity to the active substance or excipients, baseline neutrophil count of <1,500 cells/mm^3, pregnancy or breast-feeding, severe liver impairment. **Precautions and Warnings:** Reduce dosage with febrile neutropenia, neutrophils < 500 cells/mm^3 for more than one week, severe or cumulative cutaneous reactions, severe peripheral neuropathy, or moderately raised LFTs, ALT and/or AST >1.5 times the ULN concurrent with serum alkaline phosphatase >2.5 times the ULN. Severe hypersensitivity reactions require immediate discontinuation. Severe cutaneous skin reactions, such as eruptions followed by desquamation, may require interruption or treatment discontinuation. Severe fluid retention such as pleural effusion, pericardial effusion or ascites should be monitored closely. With serum bilirubin levels > ULN and/or ALT and AST >3.5 times the ULN concurrent with alkaline phosphatase levels > 6 times the ULN, no dose-reduction can be recommended and docetaxel should not be used unless strictly indicated. **Interactions:** Caution with compounds that induce, inhibit or are metabolised by cytochrome P450-3A, which may alter docetaxel metabolism. **Pregnancy and Lactation:** Contraindicated. **Adverse Reactions:** Neutropenia, thrombocytopenia, anaemia, hypersensitivity reactions, fluid retention, cutaneous reactions, arthralgia, myalgia, peripheral neuropathy and rarely other neurological events, infectious episodes, increased liver enzyme levels, alopecia, asthenia, mucositis, injection site reactions, gastrointestinal events, cardiovascular events (including hypotension and dysrhythmia). Acute respiratory distress syndrome, radiation recall phenomena, lacrimation without conjunctivitis and lacrimal duct obstruction have been rarely reported. **Pharmaceutical Precautions:** Store vials between +2°C and +25°C; protect from bright light. Reconstitute concentrate with accompanying solvent and dilute with infusion solution (0.9% sodium chloride or 5% dextrose for intravenous injection) before use. Apply usual cytotoxic precautions. **Package Quantities and Basic NHS Price:** Blister cartons containing one vial of TAXOTERE® concentrate and one vial of solvent: TAXOTERE® 20mg £175.00; TAXOTERE® 80mg £575.00. **Legal Category:** POM.
Marketing Authorisation Numbers: TAXOTERE® 20mg EU/1/95/002/001; TAXOTERE® 80mg EU/1/95/002/002. Further information available on request from Aventis Pharma Ltd., 50 Kings Hill Avenue, Kings Hill, West Malling, Kent ME19 4AH.
Date of Revision: May 2002.

Sponsored by an educational grant from
Aventis Pharma

Published by Hayward Medical Communications, a division of Hayward Group plc.
'What is ...' is a Hayward Group plc publication.

HE081002
Date of preparation: November 2002

International Journal for Quality in Health Care 2003; Volume 5, Number 3: pp. 26 –266 0. 093/intqhc/mzg03

Methodology Matters

Good practice in the conduct and reporting of survey research

KATE KELLEY, BELINDA CLARK, VIVIENNE BROWN AND JOHN SITZIA

Research Department, Worthing & Southlands Hospitals NHS Trust, Worthing, West Sussex, UK

Abstract

Survey research is sometimes regarded as an easy research approach. However, as with any other research approach and method, it is easy to conduct a survey of poor quality rather than one of high quality and real value. This paper provides a checklist of good practice in the conduct and reporting of survey research. Its purpose is to assist the novice researcher to produce survey work to a high standard, meaning a standard at which the results will be regarded as credible. The paper first provides an overview of the approach and then guides the reader step-by-step through the processes of data collection, data analysis, and reporting. It is not intended to provide a manual of how to conduct a survey, but rather to identify common pitfalls and oversights to be avoided by researchers if their work is to be valid and credible.

Keywords: data reporting, health care surveys, methodology, questionnaires, research design, survey methods, surveys

What is survey research?

Survey research is common in studies of health and health services, although its roots lie in the social surveys conducted in Victorian Britain by social reformers to collect information on poverty and working class life (e.g. Charles Booth [1] and Joseph Rowntree [2]), and indeed survey research remains most used in applied social research. The term 'survey' is used in a variety of ways, but generally refers to the selection of a relatively large sample of people from a pre-determined population (the 'population of interest'; this is the wider group of people in whom the researcher is interested in a particular study), followed by the collection of a relatively small amount of data from those individuals. The researcher therefore uses information from a sample of individuals to make some inference about the wider population.

Data are collected in a standardized form. This is usually, but not necessarily, done by means of a questionnaire or interview. Surveys are designed to provide a 'snapshot of how things are at a specific time' [3]. There is no attempt to control conditions or manipulate variables; surveys do not allocate participants into groups or vary the treatment they receive. Surveys are well suited to descriptive studies, but can also be used to explore aspects of a situation, or to seek explanation and provide data for testing hypotheses. It is important to recognize that 'the survey approach is a research strategy, not a research method' [3]. As with any research approach, a choice of methods is available and the one most appropriate to the individual project should be used. This paper will discuss the most popular methods employed in survey research, with an emphasis upon difficulties commonly encountered when using these methods.

Descriptive research

Descriptive research is a most basic type of enquiry that aims to observe (gather information on) certain phenomena, typically at a single point in time: the 'cross-sectional' survey. The aim is to examine a situation by describing important factors associated with that situation, such as demographic, socio-economic, and health characteristics, events, behaviours, attitudes, experiences, and knowledge. Descriptive studies are used to estimate specific parameters in a population (e.g. the prevalence of infant breast feeding) and to describe associations (e.g. the association between infant breast feeding and maternal age).

Analytical studies

Analytical studies go beyond simple description; their intention is to illuminate a specific problem through focused

Address reprint requests to John Sitzia, Research Department, Worthing Hospital, Lyndhurst Road, Worthing BN 2DH, West Sussex, UK. E-mail: john.sitzia@wash.nhs.uk

data analysis, typically by looking at the effect of one set of variables upon another set. These are longitudinal studies, in which data are collected at more than one point in time with the aim of illuminating the direction of observed associations. Data may be collected from the same sample on each occasion (cohort or panel studies) or from a different sample at each point in time (trend studies).

Evaluation research

This form of research collects data to ascertain the effects of a planned change.

Advantages and disadvantages of survey research

Advantages:

- The research produces data based on real-world observations (empirical data).
- The breadth of coverage of many people or events means that it is more likely than some other approaches to obtain data based on a representative sample, and can therefore be generalizable to a population.
- Surveys can produce a large amount of data in a short time for a fairly low cost. Researchers can therefore set a finite time-span for a project, which can assist in planning and delivering end results.

Disadvantages:

- The significance of the data can become neglected if the researcher focuses too much on the range of coverage to the exclusion of an adequate account of the implications of those data for relevant issues, problems, or theories.
- The data that are produced are likely to lack details or depth on the topic being investigated.
- Securing a high response rate to a survey can be hard to control, particularly when it is carried out by post, but is also difficult when the survey is carried out face-to-face or over the telephone.

Essential steps in survey research

Research question

Good research has the characteristic that its purpose is to address a single clear and explicit research question; conversely, the end product of a study that aims to answer a number of diverse questions is often weak. Weakest of all, however, are those studies that have no research question at all and whose design simply is to collect a wide range of data and then to 'trawl' the data looking for 'interesting' or 'significant' associations. This is a trap novice researchers in particular fall into. Therefore, in developing a research question, the following aspects should be considered [4]:

- Be knowledgeable about the area you wish to research.
- Widen the base of your experience, explore related areas, and talk to other researchers and practitioners in the field you are surveying.
- Consider using techniques for enhancing creativity, for example brainstorming ideas.
- Avoid the pitfalls of: allowing a decision regarding methods to decide the questions to be asked; posing research questions that cannot be answered; asking questions that have already been answered satisfactorily.

Research methods

The survey approach can employ a range of methods to answer the research question. Common survey methods include postal questionnaires, face-to-face interviews, and telephone interviews.

Postal questionnaires

This method involves sending questionnaires to a large sample of people covering a wide geographical area. Postal questionnaires are usually received 'cold', without any previous contact between researcher and respondent. The response rate for this type of method is usually low, $\sim 20\%$, depending on the content and length of the questionnaire. As response rates are low, a large sample is required when using postal questionnaires, for two main reasons: first, to ensure that the demographic profile of survey respondents reflects that of the survey population; and secondly, to provide a sufficiently large data set for analysis.

Face-to-face interviews

Face-to-face interviews involve the researcher approaching respondents personally, either in the street or by calling at people's homes. The researcher then asks the respondent a series of questions and notes their responses. The response rate is often higher than that of postal questionnaires as the researcher has the opportunity to sell the research to a potential respondent. Face-to-face interviewing is a more costly and time-consuming method than the postal survey, however the researcher can select the sample of respondents in order to balance the demographic profile of the sample.

Telephone interviews

Telephone surveys, like face-to-face interviews, allow a two-way interaction between researcher and respondent. Telephone surveys are quicker and cheaper than face-to-face interviewing. Whilst resulting in a higher response rate than postal surveys, telephone surveys often attract a higher level of refusals than face-to-face interviews as people feel less inhibited about refusing to take part when approached over the telephone.

Designing the research tool

Whether using a postal questionnaire or interview method, the questions asked have to be carefully planned and piloted.

The design, wording, form, and order of questions can affect the type of responses obtained, and careful design is needed to minimize bias in results. When designing a questionnaire or question route for interviewing, the following issues should be considered: (1) planning the content of a research tool; (2) questionnaire layout; (3) interview questions; (4) piloting; and (5) covering letter.

Planning the content of a research tool

The topics of interest should be carefully planned and relate clearly to the research question. It is often useful to involve experts in the field, colleagues, and members of the target population in question design in order to ensure the validity of the coverage of questions included in the tool (content validity).

Researchers should conduct a literature search to identify existing, psychometrically tested questionnaires. A well designed research tool is simple, appropriate for the intended use, acceptable to respondents, and should include a clear and interpretable scoring system. A research tool must also demonstrate the psychometric properties of reliability (consistency from one measurement to the next), validity (accurate measurement of the concept), and, if a longitudinal study, responsiveness to change [5]. The development of research tools, such as attitude scales, is a lengthy and costly process. It is important that researchers recognize that the development of the research tool is equal in importance—and deserves equal attention—to data collection. If a research instrument has not undergone a robust process of development and testing, the credibility of the research findings themselves may legitimately be called into question and may even be completely disregarded. Surveys of patient satisfaction and similar are commonly weak in this respect; one review found that only 6% of patient satisfaction studies used an instrument that had undergone even rudimentary testing [6]. Researchers who are unable or unwilling to undertake this process are strongly advised to consider adopting an existing, robust research tool.

Questionnaire layout

Questionnaires used in survey research should be clear and well presented. The use of capital (upper case) letters only should be avoided, as this format is hard to read. Questions should be numbered and clearly grouped by subject. Clear instructions should be given and headings included to make the questionnaire easier to follow.

The researcher must think about the form of the questions, avoiding 'double-barrelled' questions (two or more questions in one, e.g. 'How satisfied were you with your personal nurse and the nurses in general?'), questions containing double negatives, and leading or ambiguous questions. Questions may be open (where the respondent composes the reply) or closed (where pre-coded response options are available, e.g. multiple-choice questions). Closed questions with pre-coded response options are most suitable for topics where the possible responses are known. Closed questions are quick to administer and can be easily coded and analysed. Open questions should be used where possible replies are unknown or too numerous to pre-code. Open questions are more demanding for respondents but if well answered can provide useful insight into a topic. Open questions, however, can be time consuming to administer and difficult to analyse. Whether using open or closed questions, researchers should plan clearly how answers will be analysed.

Interview questions

Open questions are used more frequently in unstructured interviews, whereas closed questions typically appear in structured interview schedules. A structured interview is like a questionnaire that is administered face to face with the respondent. When designing the questions for a structured interview, the researcher should consider the points highlighted above regarding questionnaires. The interviewer should have a standardized list of questions, each respondent being asked the same questions in the same order. If closed questions are used the interviewer should also have a range of pre-coded responses available.

If carrying out a semi-structured interview, the researcher should have a clear, well thought out set of questions; however, the questions may take an open form and the researcher may vary the order in which topics are considered.

Piloting

A research tool should be tested on a pilot sample of members of the target population. This process will allow the researcher to identify whether respondents understand the questions and instructions, and whether the meaning of questions is the same for all respondents. Where closed questions are used, piloting will highlight whether sufficient response categories are available, and whether any questions are systematically missed by respondents.

When conducting a pilot, the same procedure as as that to be used in the main survey should be followed; this will highlight potential problems such as poor response.

Covering letter

All participants should be given a covering letter including information such as the organization behind the study, including the contact name and address of the researcher, details of how and why the respondent was selected, the aims of the study, any potential benefits or harm resulting from the study, and what will happen to the information provided. The covering letter should both encourage the respondent to participate in the study and also meet the requirements of informed consent (see below).

Sample and sampling

The concept of sample is intrinsic to survey research. Usually, it is impractical and uneconomical to collect data from every single person in a given population; a sample of the population has to be selected [7]. This is illustrated in the following

hypothetical example. A hospital wants to conduct a satisfaction survey of the 1000 patients discharged in the previous month; however, as it is too costly to survey each patient, a sample has to be selected. In this example, the researcher will have a list of the population members to be surveyed (sampling frame). It is important to ensure that this list is both up-to date and has been obtained from a reliable source.

The method by which the sample is selected from a sampling frame is integral to the external validity of a survey: the sample has to be representative of the larger population to obtain a composite profile of that population [8].

There are methodological factors to consider when deciding who will be in a sample: How will the sample be selected? What is the optimal sample size to minimize sampling error? How can response rates be maximized?

The survey methods discussed below influence how a sample is selected and the size of the sample. There are two categories of sampling: random and non-random sampling, with a number of sampling selection techniques contained within the two categories. The principal techniques are described here [9].

Random sampling

Generally, random sampling is employed when quantitative methods are used to collect data (e.g. questionnaires). Random sampling allows the results to be generalized to the larger population and statistical analysis performed if appropriate. The most stringent technique is simple random sampling. Using this technique, each individual within the chosen population is selected by chance and is equally as likely to be picked as anyone else. Referring back to the hypothetical example, each patient is given a serial identifier and then an appropriate number of the 1000 population members are randomly selected. This is best done using a random number table, which can be generated using computer software (a free on-line randomizer can be found at http://www.randomizer.org/index.htm).

Alternative random sampling techniques are briefly described. In systematic sampling, individuals to be included in the sample are chosen at equal intervals from the population; using the earlier example, every fifth patient discharged from hospital would be included in the survey. Stratified sampling selects a specific group and then a random sample is selected. Using our example, the hospital may decide only to survey older surgical patients. Bigger surveys may employ cluster sampling, which randomly assigns groups from a large population and then surveys everyone within the groups, a technique often used in national-scale studies.

Non-random sampling

Non-random sampling is commonly applied when qualitative methods (e.g. focus groups and interviews) are used to collect data, and is typically used for exploratory work. Non-random sampling deliberately targets individuals within a population. There are three main techniques. (1) purposive sampling: a specific population is identified and only its members are included in the survey; using our example above, the hospital may decide to survey only patients who had an appendectomy. (2) Convenience sampling: the sample is made up of the individuals who are the easiest to recruit. Finally, (3) snowballing: the sample is identified as the survey progresses; as one individual is surveyed he or she is invited to recommend others to be surveyed.

It is important to use the right method of sampling and to be aware of the limitations and statistical implications of each. The need to ensure that the sample is representative of the larger population was highlighted earlier and, alongside the sampling method, the degree of sampling error should be considered. Sampling error is the probability that any one sample is not completely representative of the population from which it has been drawn [9]. Although sampling error cannot be eliminated entirely, the sampling technique chosen will influence the extent of the error. Simple random sampling will give a closer estimate of the population than a convenience sample of individuals who just happened to be in the right place at the right time.

Sample size

What sample size is required for a survey? There is no definitive answer to this question: large samples with rigorous selection are more powerful as they will yield more accurate results, but data collection and analysis will be proportionately more time consuming and expensive. Essentially, the target sample size for a survey depends on three main factors: the resources available, the aim of the study, and the statistical quality needed for the survey. For 'qualitative' surveys using focus groups or interviews, the sample size needed will be smaller than if quantitative data is collected by questionnaire. If statistical analysis is to be performed on the data then sample size calculations should be conducted. This can be done using computer packages such as G*Power [10]; however, those with little statistical knowledge should consult a statistician. For practical recommendations on sample size, the set of survey guidelines developed by the UK Department of Health [11] should be consulted.

Larger samples give a better estimate of the population but it can be difficult to obtain an adequate number of responses. It is rare that everyone asked to participate in the survey will reply. To ensure a sufficient number of responses, include an estimated non-response rate in the sample size calculations.

Response rates are a potential source of bias. The results from a survey with a large non-response rate could be misleading and only representative of those who replied. French [12] reported that non-responders to patient satisfaction surveys are less likely to be satisfied than people who reply. It is unwise to define a level above which a response rate is acceptable, as this depends on many local factors; however, an achievable and acceptable rate is ~ 75% for interviews and 65% for self-completion postal questionnaires [9,13]. In any study, the final response rate should be reported with the results; potential differences between the respondents and non-respondents should be explicitly explored and their implications discussed.

There are techniques to increase response rates. A questionnaire must be concise and easy to understand, reminders should be sent out, and method of recruitment should be carefully considered. Sitzia and Wood [13] found that participants recruited by mail or who had to respond by mail had a lower mean response rate (67%) than participants who were recruited personally (mean response 76.7%). A most useful review of methods to maximize response rates in postal surveys has recently been published [14].

Data collection

Researchers should approach data collection in a rigorous and ethical manner. The following information must be clearly recorded:

- How, where, how many times, and by whom potential respondents were contacted.
- How many people were approached and how many of those agreed to participate.
- How did those who agreed to participate differ from those who refused with regard to characteristics of interest in the study, for example how were they identified, where were they approached, and what was their gender, age, and features of their illness or health care.
- How was the survey administered (e.g. telephone interview).
- What was the response rate (i.e. the number of usable data sets as a proportion of the number of people approached).

Data analysis

The purpose of all analyses is to summarize data so that it is easily understood and provides the answers to our original questions: 'In order to do this researchers must carefully examine their data; they should become friends with their data' [15]. Researchers must prepare to spend substantial time on the data analysis phase of a survey (and this should be built into the project plan). When analysis is rushed, often important aspects of the data are missed and sometimes the wrong analyses are conducted, leading to both inaccurate results and misleading conclusions [16]. However, and this point cannot be stressed strongly enough, researchers must not engage in data dredging, a practice that can arise especially in studies in which large numbers of dependent variables can be related to large numbers of independent variables (outcomes). When large numbers of possible associations in a dataset are reviewed at $P < 0.05$, one in 20 of the associations by chance will appear 'statistically significant'; in datasets where only a few real associations exist, testing at this significance level will result in the large majority of findings still being false positives [17].

The method of data analysis will depend on the design of the survey and should have been carefully considered in the planning stages of the survey. Data collected by qualitative methods should be analysed using established methods such as content analysis [18], and where quantitative methods have been used appropriate statistical tests can be applied. Describing methods of analysis here would be unproductive as a multitude of introductory textbooks and on-line resources are available to help with simple analyses of data (e.g. [19, 20]). For advanced analysis a statistician should be consulted.

Reporting

When reporting survey research, it is essential that a number of key points are covered (though the length and depth of reporting will be dependent upon journal style). These key points are presented as a 'checklist' below:

(1) Explain the purpose or aim of the research, with the explicit identification of the research question.
(2) Explain why the research was necessary and place the study in context, drawing upon previous work in relevant fields (the literature review).
(3) Describe in (proportionate) detail how the research was done.
    (a) State the chosen research method or methods, and justify why this method was chosen.
    (b) Describe the research tool. If an existing tool is used, briefly state its psychometric properties and provide references to the original development work. If a new tool is used, you should include an entire section describing the steps undertaken to develop and test the tool, including results of psychometric testing.
    (c) Describe how the sample was selected and how data were collected, including:
        (i) How were potential subjects identified?
        (ii) How many and what type of attempts were made to contact subjects?
        (iii) Who approached potential subjects?
        (iv) Where were potential subjects approached?
        (v) How was informed consent obtained?
        (vi) How many agreed to participate?
        (vii) How did those who agreed differ from those who did not agree?
        (viii) What was the response rate?
(4) Describe and justify the methods and tests used for data analysis.
(5) Present the results of the research. The results section should be clear, factual, and concise.
(6) Interpret and discuss the findings. This 'discussion' section should not simply reiterate results; it should provide the author's critical reflection upon both the results and the processes of data collection. The discussion should assess how well the study met the research question, should describe the problems encountered in the research, and should honestly judge the limitations of the work.
(7) Present conclusions and recommendations.

The researcher needs to tailor the research report to meet:

- The expectations of the specific audience for whom the work is being written.
- The conventions that operate at a general level with respect to the production of reports on research in the social sciences.

Ethics

Anyone involved in collecting data from patients has an ethical duty to respect each individual participant's autonomy. Any survey should be conducted in an ethical manner and one that accords with best research practice. Two important ethical issues to adhere to when conducting a survey are confidentiality and informed consent.

The respondent's right to confidentiality should always be respected and any legal requirements on data protection adhered to. In the majority of surveys, the patient should be fully informed about the aims of the survey, and the patient's consent to participate in the survey must be obtained and recorded.

The professional bodies listed below, among many others, provide guidance on the ethical conduct of research and surveys.

- American Psychological Association: http://www.apa.org
- British Psychological Society: http://www.bps.org.uk
- British Medical Association: http://www.bma.org.uk.
- UK General Medical Council: http://www.gmc-uk.org
- American Medical Association: http://www.ama-assn.org
- UK Royal College of Nursing: http://www.rcn.org.uk
- UK Department of Health: http://www.doh.gov

Conclusion

Survey research demands the same standards in research practice as any other research approach, and journal editors and the broader research community will judge a report of survey research with the same level of rigour as any other research report. This is not to say that survey research need be particularly difficult or complex; the point to emphasize is that researchers should be aware of the steps required in survey research, and should be systematic and thoughtful in the planning, execution, and reporting of the project. Above all, survey research should not be seen as an easy, 'quick and dirty' option; such work may adequately fulfil local needs (e.g. a quick survey of hospital staff satisfaction), but will not stand up to academic scrutiny and will not be regarded as having much value as a contribution to knowledge.

References

1. London School of Economics, UK. Http://booth.lse.ac.uk/ (accessed 15 January 2003).
2. Vernon A. A Quaker Businessman: Biography of Joseph Rowntree (1836□1925). London: Allen & Unwin, 1958.
3. Denscombe M. The Good Research Guide: For Small-scale Social Research Projects. Buckingham: Open University Press, 1998.
4. Robson C. Real World Research: A Resource for Social Scientists and Practitioner-researchers. Oxford: Blackwell Publishers, 1993.
5. Streiner DL, Norman GR. Health Measurement Scales: A Practical Guide to their Development and Use. Oxford: Oxford University Press, 1995.
6. Sitzia J. How valid and reliable are patient satisfaction data? An analysis of 195 studies. Int J Qual Health Care 1999; 11: 319–328.
7. Bowling A. Research Methods in Health. Investigating Health and Health Services. Buckingham: Open University Press, 2002.
8. American Statistical Association, USA. Http://www.amstat.org (accessed 9 December 2002).
9. Arber S. Designing samples. In: Gilbert N, ed. Researching Social Life. London: SAGE Publications, 2001.
10. Heinrich Heine University, Dusseldorf, Germany. Http://www.psycho.uni-duesseldorf.de/aap/projects/gpower/index.html (accessed 12 December 2002).
11. Department of Health, England. Http://www.doh.gov.uk/acutesurvey/index.htm (accessed 12 December 2002).
12. French K. Methodological considerations in hospital patient opinion surveys. Int J Nurs Stud 1981; 18: 7–32.
13. Sitzia J, Wood N. Response rate in patient satisfaction research: an analysis of 210 published studies. Int J Qual Health Care 1998; 10: 311–317.
14. Edwards P, Roberts I, Clarke M et al. Increasing response rates to postal questionnaires: systematic review. Br Med J 2002; 324: 1183.
15. Wright DB. Making friends with our data: improving how statistical results are reported. Br J Educ Psychol 2003; in press.
16. Wright DB, Kelley K. Analysing and reporting data. In: Michie S, Abraham C, eds. Health Psychology in Practice. London: SAGE Publications, 2003; in press.
17. Davey Smith G, Ebrahim S. Data dredging, bias, or confounding. Br Med J 2002; 325: 1437–1438.
18. Morse JM, Field PA. Nursing Research: The Application of Qualitative Approaches. London: Chapman and Hall, 1996.
19. Wright DB. Understanding Statistics: An Introduction for the Social Sciences. London: SAGE Publications, 1997.
20. Sportscience, New Zealand. Http://www.sportsci.org/resource/stats/index.html (accessed 12 December 2002).

Accepted for publication 16 January 2003

Styles of Supervision

The Meeting

Styles of Supervision

High Support	**Pastoral Style** **Low structure and high support** • Student has personal low management skill but takes advantage of all the support facilities that are on offer • Supervisor provides considerable personal care and support but not necessarily in a task-driven, directive capacity	**Contractual Style** **High structure and high support** • Student highly motivated and able to take direction and to act on own initiative • Supervisor able to administer direction and exercises good management skills and interpersonal relationships
Low Support	**Laissez-faire Style** **Low structure low support** • Student has limited levels of motivation and management skills • Supervisor in non-directive and not committed to high levels of personal interaction Supervisor may appear uncaring and uninvolved **Low Structure**	**Directorial style** **High structure and low support** • Student highly motivated and sees the necessity to take advantage of engaging in high structural activities such as setting objectives, completing and submitting work o time on own initiative without taking advantage of institutional support. • Supervisor has a close an d regular interactive relationship with the candidate, but avoids non-task issues. **High structure**

In pairs, discuss how you would adjust your supervisory style with each of the following students:

1. Susan is a single mother holding down a part-time job. She is usually harried and stressed because of the multiple tasks she is trying to juggle and doubtful of her own capabilities of completing the PhD.
2. Yusuf is a devout 54 year-old Muslim who decided to be the first in his family to do a Master's. He is anxious not to let his family down. He is extremely deferential to the supervisor and unsure of his own ideas.
3. Dumisa is a young, very smart, accomplished student who has ambitions to have an academic career. She graduated *cum laude* on her Master's.
4. Biniam is an international student from Ethiopia on a 3 year scholarship to complete his PhD. Although academically astute, he struggles with the language and adjusting to the culture at UCT.

RESEARCH STUDENT SUPERVISORY MEETING CHECKLIST

INITIAL MEETING

Degree:
Date:
Student:
Supervisor 1: _____ Supervisor 2: _____

Meeting Tasks	Confirmed
Student/registrar roles and responsibilities discussed	
Supervisor 1 roles and responsibilities discussed	
*Supervisor 2 roles and responsibilities discussed	
Contact arrangements agreed	
Frequency of meetings agreed	
Feedback: mode and timing agreed	
MOU signed	
Research question and design discussed	
Sample size	
Research time allocation	
Student/registrar guided to relevant literature	
Training needs identified e.g. database search, stats, design, clinical techs	
Necessary permissions discussed and contacts provided where possible	
Progress of proposal, ethical, legal and safety issues discussed	
Dates for forms and registrations at postgrad offices	
Student/registrar informed of UCT wide seminars/training	
Next milestone identified and next meeting date set	

*Registrars may only have one supervisor.

RESEARCH STUDENT SUPERVISORY MEETING CHECKLIST

SUBSEQUENT MEETINGS

Degree:
Date:
Student:
Supervisor 1: _____ Supervisor 2: _____

Meeting Tasks	**Confirmed**
Agenda sent by student/registrar in advance	
Team prepared	
Review	
Previous action points Protocol completion/ethics submission done	
Definition	
Scope and purpose of current meeting clear	
Exploration	
Literature review discussed and EQUATOR and Critical appraisal websites provided	
Results discussed	
Intra and interdisciplinary contacts provided where necessary	
Progress monitored: timeline/milestones/scope	
Clarification	
Meeting summarized and decisions are made by student/registrar	
Goal setting	
Realistic action plan agreed	
Presentation/publication and IP opportunities discussed	
Next milestone identified and next meeting date set	

Styles of supervision

UCT CLINICAL RESEARCH CENTRE
every step of the study

Clinical Research Supervision

ACTIVITY 1

1.Using the styles of supervision rubric , select the one that most closely reflects your choice.

2. Discuss with the person next to you why you have chosen that style, eg. It suits your lifestyle or your personality.

GROUP DISCUSSION

BARRIERS TO EFFECTIVE SUPERVISION

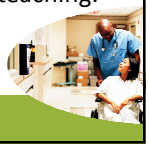
- Increasing numbers of registrars.
- Limited funds for new academic staff
- Limited funds for MMed research resources eg. Statistics
- Increase in workload for supervisors and registrars threatens quality of research.

WHAT IS EFFECTIVE SUPERVISION

THE SUPERVISORY MEETING

COMMON MISCONCEPTIONS

- Registrar should be able to work independently.
- Corridor/brief meetings are sufficient.
- Supervision of research is not about teaching.



The Aims of Meetings

Level of study

Registrar level at entry

University guidelines

Project requirements

Level of Study: Descriptors

MMED – ability to:

- undertake research
- interpret results adequately
- review the relevant literature comprehensively and **critically**.

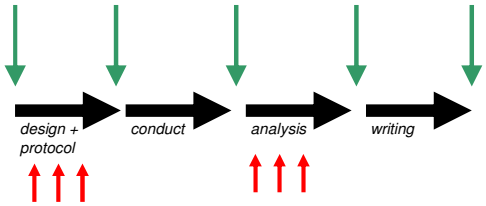
Need not be original

Frequency & length of meetings

- Should be based on structure (ad hoc never benefits anyone; appeals)
- Should meet student agenda.
- Should be dedicated time

Frequency of meetings –MMED

at least **5 doses of substantial support** from supervisor/s



There are a few points where ***specific forms of assistance*** *are helpful*

All meetings must consider

MMed Process

- UCT Milestones
- Study Milestones
- Holiday/travel cover
- Training needs
- Resource access

Academic Development

- Adequate review of any material submitted by student
- Intellectual debate
- Literature discussions

UCT CLINICAL RESEARCH CENTRE

MEETING 1:PROCESS

- Informal & registrar takes mins & creates action points
- Clarify expectations: Roles and responsibilities discussed & agreed & signed.
- MOU
- Feedback – how much, how soon, format
- Next date set.

UCT CLINICAL RESEARCH CENTRE

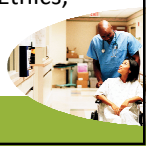
MEETING 1:ACADEMIC

- Review of research question & scope of project.
- Discuss & educate how research question is linked to design
- Identify required training needs.
- Direct to relevant literature.
- Discussion of contacts required to pursue project eg. access to clinics, medical records in different units and permissions reqd.

UCT CLINICAL RESEARCH CENTRE

ACADEMIC: MEETING 1 contd

- Consider possibility of IP.
- Encourage to take part in intellectual life of university.
- Educate and Guide on processes eg. Ethics, legal, safety, permissions etc



PROCESS: Subsequent Meetings

- Registrar sends agenda & team is prepared
- Review: Previous action points.
- Definition: Scope & purpose of current meeting
- Exploration: Scholarly development: results, literature
  - Progress: timeline/milestones
- Clarification: Decisions needed
- Goal setting: Decisions taken-action plan (realistic)
- Conclusion: Date of next meeting set

ACADEMIC: Subsequent Meetings

- Discussion of findings, new literature, submitted material.
- Enable student to provide educated appraisal of lit.- critical appraisal tools
- Guide student to standards and tools eg. validated outcome measures, EQUATOR
- Discussion/education on design/write-up and required training needs.

ACADEMIC: Subsequent Meetings

- Equipment and other facilities identified for OM
- Planning of presentations to dept, conf etc
- Discussion and training in how to write articles/prepare presentations

PLAN HOLIDAYS

Other important issues to be included in meetings

- Writing thesis vs publishing article
- How many publications a thesis might render.
- Authorship in joint publications.
- Process for unresolved issues of conduct.

The Write Up

SUPPORTING THE WRITE UP

UCT CLINICAL RESEARCH CENTRE
every step of the study

Clinical Research Supervision

TIMEFRAME AND SCOPE

UCT CLINICAL RESEARCH CENTRE

Training the Writer

- Starts early with proposal/ethics.
- Continuous process with regular feedback.
- Supervision has to address the process and the content.

Daunting Process !

UCT CLINICAL RESEARCH CENTRE

Why does writing scare people ?

- Making ideas explicit and putting them out there for others to review
  - Fears about their ideas being judged
- Concerns about how to express ideas on paper
  - Fears about them being judged

FEEDBACK MUST BE DELIVERED IN A COLLEGIATE AND CONSTRUCTIVE MANNER

Tell them the Good News

- This is scientific writing
  - Not creative writing
- Emphasis on being **clear & brief**
  - Provide necessary information in accessible manner
  - Short simple sentences
  - Link paragraphs
- There are guidelines
- Adhere to standard formats

General suggestions for registrar

1. Read previous MMeds (library)
2. When reading journal articles
   - Read content
   - Notice style and structure
   - If you think it is well written ask yourself why

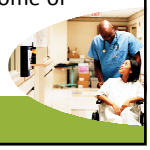
Do what they do

- Copy **structure** of sentences, paragraphs, tables, figures

General suggestions for registrar

- Work at set times and in a dedicated space.
- Start using a reference manager immediately.
- Record all thoughts, ideas, questions.
- Record all databases searched and the outcome of searches.

Invite comments.



General suggestions for registrars

– Place research within a theoretical model where possible.

– Know your reader.

ie. A reader who may know nothing about your topic, but who is capable of understanding

THE ART OF SCIENCE IN WRITING

Introduction/Lit Rev
Organise the Evidence

FIRST:

Organise your papers

THEN:

Organise your thoughts

Literature review

1. Start by writing summary on each of themes/designs/opinions – then arrange them logically, then link
2. Don't present an annotated bibliography – 'sound bytes'
   - **Don't go study by study**
     - Boring to read
     - Does not show insight
3. Summarise state of knowledge on a topic
   - Highlight key/seminal studies

Getting Started: The Introduction

1. Start with broad concepts ~ background
   - Epidemiology of condition
   - History of condition, management practices

- Narrow down to your specific question
  - Issues specific to your question
- End with your research question
  - Aims & objectives

Introduction/Lit review

KEY POINTS 1 & 2

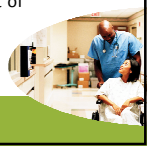
Penultimate paragraph – IDENTIFY GAP in evidence

Why are some questions still not answered?
appraise/critique/give reasons eg. design

Ultimate paragraph- therefore this study aims to .....ie. how will you fill gap/contribute to evidence.

- ***state aims as primary/secondary*** – broader statement of intent.

-state objectives for each aim – specific statement that can be measured



UCT CLINICAL RESEARCH CENTRE

KEY POINT 3

LINK SECTIONS

INTRO **M&M** **RESULTS** **DISC**

Aim

Obj 1 → How meas → outcome → Relevance

Obj 2 → How meas → outcome → Relevance

Methods – How will you measure objectives ?

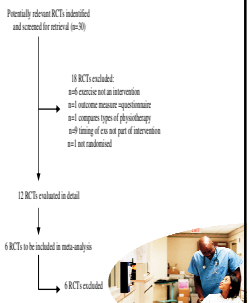
– Use published guidelines as a guide

- The 'cookbook' ~ provides a recipe for repeating the study
  – Include key documents in appendix:
  Questionnaires, data abstraction forms
- Wherever possible use validated questionnaires, measures.

UCT CLINICAL RESEARCH CENTRE

Methods: Diagrams/Tables

- Diagrams to show:
  - Study flow:
    - How participants recruited
    - How participants managed
    - Points of exclusion
- Database searches



UCT CLINICAL RESEARCH CENTRE

Variables and Observation points

Outcomes	0mo	6mo	12mo
Med history	X		
Medications	X	X	X
Disease screen	X		
Inclusion criteria	X		
Questionnaires	X	X	X
Blood tests	X		X
BP, Heart rate, vibration sensitivity	X		X
Treatment	X	X	X
Treatment outcome	X	X	X

UCT CLINICAL RESEARCH CENTRE

KEY POINT 3

LINK SECTIONS

INTRO **M&M** **RESULTS** **DISC**

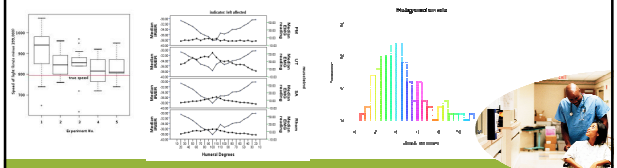
Aim

Obj 1 → How meas → outcome → Relevance

Obj 2 → How meas → outcome → Relevance

Presenting Results

- Report only their results
- Look at tables presented in relevant journal articles in your field
- Encourage use of varied forms of presentation.
- Report key findings from graphs in writing.



Guiding the Discussion

Don't cut corners here

- Well-written discussion shows an insightful student ***regardless of the quality of the data***

1. **Summarise** key findings from results: This study has shown that….

2. **Interpret the results** presented
   - Compare/contrast to existing literature
   - What is novel
   - What is similar
   - What is different
   - Why are findings similar or different?

Guiding the Discussion

3. Discuss the **strengths and limitations**

Why is your study important

- What the caveats are for interpreting data
- Be honest
- Show that you understand research methods

Studies with limitations are OK, as long as you recognise the limitations

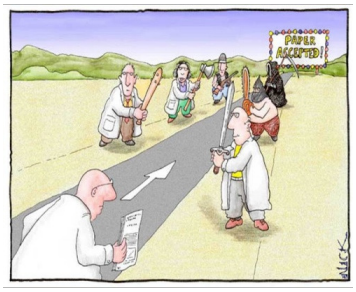
Don't pretend you have all the answers!!

Guiding the Discussion

4. Discuss the **implications of results and make recommendations**:
   - For future research
     What more studies do you think are needed?
   - For clinical services
     What would you recommend for services?
   - For policy-making and programmes
     What changes do you recommend?

Show that you have a sense of the value of research

Brave a review…….



THE ART OF LANGUAGE IN WRITING

Paragraph Structure

Component	Function
Topic sentence(s)	•To introduce and clearly state the main idea/point that you intend to develop
Development or elaboration of ideas	•To elaborate the new idea or point that you have introduced.
Concluding sentence(s)	•To round off what you have said so far in your paragraph •To link the current paragraph to the next paragraph

UCT CLINICAL RESEARCH CENTRE

Cohesion ('united whole') in paragraphs

1. Connectors – make a text 'hang together' or increase a text's **cohesion**

'hooks' that connect different parts of a text

sometimes explicitly describe the structure or development of the text e.g. ***in the first place, therefore, however, in addition…….***

Cohesion

2. Repetition of key words/terms forms '**strands**' of meaning throughout the text

3. Pronouns like **it, them, those** point to antecedents, creating links

Coherence

- Coherence refers to the overall sense or logical interconnection i.e. the pattern it follows:

a comparison

an argument

a description in time

Coherence

- Coherence includes the unity created through cohesion – particular 'patterns' or logical structures are created through connectors,

Examples:

connectors signalling comparison: **like, unlike, similarly, on the other hand...**

connectors signalling argument: ...?

Coherence

Coherence works within and across paragraphs (paragraphs are the 'building blocks' of logic)

Prospective studies have shown increased acquisition of HIV infection in uncircumcised compared with circumcised men in selected high-risk populations [1-6]. Also, in a study of discordant couples, we found lower levels of HIV acquisition in circumcised HIV-negative men [7]. However, no prospective studies have been conducted in representative general populations. Cross-sectional and ecological studies also suggest that circumcision may protect men from prevalent HIV and STD infections [8,9], and the protective effects of circumcision are most marked if the procedure is performed before the onset of puberty [10]. On the basis of these findings from observational studies, it has been proposed that circumcision should be widely promoted as a means of HIV prevention [8,9]. To assess the role of male circumcision in HIV prevention, we examined the effects of circumcision on HIV acquisition in a representative population-based cohort of……

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies

Erik von Elm, Douglas G Altman, Matthias Egger, Stuart J Pocock, Peter C Gøtzsche, Jan P Vandenbroucke, for the STROBE initiative

Much biomedical research is observational. The reporting of such research is often inadequate, which hampers the assessment of its strengths and weaknesses and of a study's generalisability. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) initiative developed recommendations on what should be included in an accurate and complete report of an observational study. We defined the scope of the recommendations to cover three main study designs: cohort, case-control, and cross-sectional studies. We convened a 2-day workshop in September, 2004, with methodologists, researchers, and journal editors to draft a checklist of items. This list was subsequently revised during several meetings of the coordinating group and in e-mail discussions with the larger group of STROBE contributors, taking into account empirical evidence and methodological considerations. The workshop and the subsequent iterative process of consultation and revision resulted in a checklist of 22 items (the STROBE statement) that relate to the title, abstract, introduction, methods, results, and discussion sections of articles. 18 items are common to all three study designs and four are specific for cohort, case-control, or cross-sectional studies. A detailed explanation and elaboration document is published separately and is freely available on the websites of *PLoS Medicine*, *Annals of Internal Medicine*, and *Epidemiology*. We hope that the STROBE statement will contribute to improving the quality of reporting of observational studies.

Lancet 2007; 370: 1453–57

Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland (E von Elm MD, Prof M Egger MD); **Centre for Statistics in Medicine, University of Oxford, Oxford, UK** (Prof D G Altman DSc); **Department of Social Medicine, University of Bristol, Bristol, UK** (M Egger); **London School of Hygiene and Tropical Medicine, University of London, London, UK** (Prof S J Pocock PhD); **Nordic Cochrane Centre, Copenhagen, Denmark** (P C Gøtzsche MD); **and Department of Clinical Epidemiology, Leiden University Hospital, Leiden, Netherlands** (Prof J P Vandenbroucke MD)

Correspondence to:
Dr Erik von Elm, Institute of Social and Preventive Medicine (ISPM), University of Bern, Finkenhubelweg 11, CH-3012 Bern, Switzerland
strobe@ispm.unibe.ch

Introduction

Many questions in medical research are investigated in observational studies.[1] Much of the research into the cause of diseases relies on cohort, case-control, or cross-sectional studies. Observational studies also have a role in research into the benefits and harms of medical interventions.[2] Randomised trials cannot answer all important questions about a given intervention. For example, observational studies are more suitable to detect rare or late adverse effects of treatments, and are more likely to provide an indication of what is achieved in daily medical practice.[3]

Research should be reported transparently so that readers can follow what was planned, what was done, what was found, and what conclusions were drawn. The credibility of research depends on a critical assessment by others of the strengths and weaknesses in study design, conduct, and analysis. Transparent reporting is also needed to judge whether and how results can be included in systematic reviews.[4,5] However, in published observational research important information is often missing or unclear. An analysis of epidemiological studies published in general medical and specialist journals found that the rationale behind the choice of potential confounding variables was often not reported.[6] Only a few reports of case-control studies in psychiatry explained the methods used to identify cases and controls.[7] In a survey of longitudinal studies in stroke research, 17 of 49 articles (35%) did not specify the eligibility criteria.[8] Others have argued that without sufficient clarity of reporting, the benefits of research might be achieved more slowly,[9] and that there is a need for guidance in reporting observational studies.[10,11]

Recommendations on the reporting of research can improve reporting quality. The Consolidated Standards of Reporting Trials (CONSORT) statement was developed in 1996 and revised 5 years later.[12] Many medical journals supported this initiative,[13] which has helped to improve the quality of reports of randomised trials.[14,15] Similar initiatives have followed for other research areas—eg, for the reporting of meta-analyses of randomised trials[16] or diagnostic studies.[17] We established a network of methodologists, researchers, and journal editors to develop recommendations for the reporting of observational research: the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement.

Aims and use of the STROBE statement

The STROBE statement is a checklist of items that should be addressed in articles reporting on the three main study designs of analytical epidemiology: cohort, case-control, and cross-sectional studies. The intention is solely to provide guidance on how to report observational research well: these recommendations are not prescriptions for designing or conducting studies. Also, while clarity of reporting is a prerequisite to evaluation, the checklist is not an instrument to evaluate the quality of observational research.

Here we present the STROBE statement and explain how it was developed. In a detailed companion paper, the explanation and elaboration article,[18–20] we justify the inclusion of the different checklist items and give methodological background and published examples of what we consider transparent reporting. We strongly recommend using the STROBE checklist in conjunction with the explanatory article, which is available freely on the websites of *PLoS Medicine* (www.plosmedicine.org), *Annals of Internal Medicine* (www.annals.org), and *Epidemiology* (www.epidem.com).

	Item	Recommendation	Reported on manuscript page
Title and abstract			
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	
Objectives	3	State specific objectives, including any prespecified hypotheses	
Methods			
Study design	4	Present key elements of study design early in the paper	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	
Participants	6	(a) *Cohort study*—give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—give the eligibility criteria, and the sources and methods of selection of participants	
		(b) *Cohort study*—for matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—for matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	
Data sources/ measurement	8*	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	
Bias	9	Describe any efforts to address potential sources of bias	
Study size	10	Explain how the study size was arrived at	
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	
		(d) *Cohort study*—if applicable, explain how loss to follow-up was addressed *Case-control study*—if applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—if applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	
Results			
Participants	13*	(a) Report the numbers of individuals at each stage of the study—eg, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg, demographic, clinical, social) and information on exposures and potential confounders	
		(b) Indicate the number of participants with missing data for each variable of interest	
		(c) *Cohort study*—summarise follow-up time (eg, average and total amount)	
Outcome data	15*	*Cohort study*—report numbers of outcome events or summary measures over time *Case-control study*—report numbers in each exposure category, or summary measures of exposure *Cross-sectional study*—report numbers of outcome events or summary measures	
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	
		(b) Report category boundaries when continuous variables were categorised	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg, analyses of subgroups and interactions, and sensitivity analyses	
Discussion			
Key results	18	Summarise key results with reference to study objectives	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	
Generalisability	21	Discuss the generalisability (external validity) of the study results	
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	

*Give such information separately for cases and controls in case-control studies, and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies. An explanation and elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the websites of *PLoS Medicine*, *Annals of Internal Medicine*, and *Epidemiology*). Separate versions of the checklist for cohort, case-control, and cross-sectional studies are available on the STROBE website.

Table: **The STROBE statement—checklist of items that should be addressed in reports of observational studies**

Development of the STROBE statement

We established the STROBE initiative in 2004, obtained funding for a workshop, and set up a website (www.strobe-statement.org). We searched textbooks, bibliographic databases, reference lists, and personal files for relevant material, including previous recommendations, empirical studies of reporting, and articles describing relevant methodological research. Because observational research makes use of many different study designs, we felt that the scope of STROBE had to be clearly defined early on. We decided to focus on the three study designs that are used most widely in analytical observational research: cohort, case-control, and cross-sectional studies.

We organised a 2-day workshop in Bristol, UK, in September, 2004. 23 individuals attended this meeting, including editorial staff from *Annals of Internal Medicine*, *BMJ*, *Bulletin of the World Health Organization*, *International Journal of Epidemiology*, *JAMA*, *Preventive Medicine*, and *The Lancet*, as well as epidemiologists, methodologists, statisticians, and practitioners from Europe and North America. Written contributions were sought from ten other individuals who declared an interest in contributing to STROBE, but could not attend. Three working groups identified items deemed to be important to include in checklists for each type of study. A provisional list of items prepared in advance (available from our website) was used to facilitate discussions. The three draft checklists were then discussed by all participants and, where possible, items were revised to make them applicable to all three study designs. In a final plenary session, the group decided on the strategy for finalising and disseminating the STROBE statement.

After the workshop we drafted a combined checklist including all three designs and made it available on our website. We invited participants and additional scientists and editors to comment on this draft checklist. We subsequently published three revisions on the website, and two summaries of comments received and changes made. During this process the coordinating group (ie, the authors of the present paper) met on eight occasions for 1 or 2 days and held several telephone conferences to revise the checklist and to prepare the present paper and the explanation and elaboration paper.[18–20] The coordinating group invited three additional co-authors with methodological and editorial expertise to help write the explanation and elaboration paper, and sought feedback from more than 30 people, who are listed at the end of this paper. We allowed several weeks for comments on subsequent drafts of the paper and reminded collaborators about deadlines by e-mail.

STROBE components

The STROBE statement is a checklist of 22 items that we consider essential for good reporting of observational studies (table). These items relate to the article's title and abstract (item 1), the introduction (items 2 and 3), methods (items 4–12), results (items 13–17), and discussion sections (items 18–21), and other information (item 22 on funding). 18 items are common to all three designs, while four (items 6, 12, 14, and 15) are design-specific, with different versions for all or part of the item. For some items (indicated by asterisks), information should be given separately for cases and controls in case-control studies, or exposed and unexposed groups in cohort and cross-sectional studies. Although presented here as a single checklist, separate checklists are available for each of the three study designs on the STROBE website.

For more on **the STROBE initiative** see www.strobe-statement.org

Implications and limitations

The STROBE statement was developed to assist authors when writing up analytical observational studies, to support editors and reviewers when considering such articles for publication, and to help readers when critically appraising published articles. We developed the checklist through an open process, taking into account the experience gained with previous initiatives, in particular CONSORT. We reviewed the relevant empirical evidence as well as methodological work, and subjected consecutive drafts to an extensive iterative process of consultation. The checklist presented here is thus based on input from a large number of individuals with diverse backgrounds and perspectives. The comprehensive explanatory article,[18–20] which is intended for use alongside the checklist, also benefited greatly from this consultation process.

Observational studies serve a wide range of purposes, on a continuum from the discovery of new findings to the confirmation or refutation of previous findings.[18–20] Some studies are essentially exploratory and raise interesting hypotheses. Others pursue clearly defined hypotheses in available data. In yet another type of studies, the collection of new data is planned carefully on the basis of an existing hypothesis. We believe the present checklist can be useful for all these studies, since the readers always need to know what was planned (and what was not), what was done, what was found, and what the results mean. We acknowledge that STROBE is currently limited to three main observational study designs. We would welcome extensions that adapt the checklist to other designs—eg, case-crossover studies or ecological studies—and also to specific topic areas. Four extensions are now available for the CONSORT statement.[21–24] A first extension to STROBE is underway for gene-disease association studies: the STROBE Extension to Genetic Association studies (STREGA) initiative.[25] We ask those who aim to develop extensions of the STROBE statement to contact the coordinating group first to avoid duplication of effort.

The STROBE statement should not be interpreted as an attempt to prescribe the reporting of observational research in a rigid format. The checklist items should be addressed in sufficient detail and with clarity somewhere in an article, but the order and format for presenting information

depends on author preferences, journal style, and the traditions of the research field. For instance, we discuss the reporting of results under a number of separate items, while recognising that authors might address several items within a single section of text or in a table. Also, item 22, on the source of funding and the role of funders, could be addressed in an appendix or in the methods section of the article. We do not aim at standardising reporting. Authors of randomised clinical trials were asked by an editor of a specialist medical journal to "CONSORT" their manuscripts on submission.[26] We believe that manuscripts should not be "STROBEd", in the sense of regulating style or terminology. We encourage authors to use narrative elements, including the description of illustrative cases, to complement the essential information about their study, and to make their articles an interesting read.[27]

We emphasise that the STROBE statement was not developed as a tool for assessing the quality of published observational research. Such instruments have been developed by other groups and were the subject of a recent systematic review.[28] In the explanation and elaboration paper, we used several examples of good reporting from studies whose results were not confirmed in further research—the important feature was the good reporting, not whether the research was of good quality. However, if STROBE is adopted by authors and journals, issues such as confounding, bias, and generalisability could become more transparent, which might help temper the over-enthusiastic reporting of new findings in the scientific community and popular media,[29] and improve the methodology of studies in the long term. Better reporting may also help to have more informed decisions about when new studies are needed, and what they should address.

We did not undertake a comprehensive systematic review for each of the checklist items and subitems, or do our own research to fill gaps in the evidence base. Further, although no one was excluded from the process, the composition of the group of contributors was influenced by existing networks and was not representative in terms of geography (it was dominated by contributors from Europe and North America) and probably was not representative in terms of research interests and disciplines. We stress that STROBE and other recommendations on the reporting of research should be seen as evolving documents that require continual assessment, refinement, and, if necessary, change. We welcome suggestions for the further dissemination of STROBE—eg, by re-publication of the present article in specialist journals and in journals published in other languages. Groups or individuals who intend to translate the checklist to other languages should consult the coordinating group beforehand. We will revise the checklist in the future, taking into account comments, criticism, new evidence, and experience from its use. We invite readers to submit their comments via the STROBE website.

Contributors

The following individuals have contributed to the content and elaboration of the STROBE statement: Douglas G Altman, Maria Blettner, Paolo Boffetta, Hermann Brenner, Geneviève Chêne, Cyrus Cooper, George Davey-Smith, Erik von Elm, Matthias Egger, France Gagnon, Peter C Gøtzsche, Philip Greenland, Sander Greenland, Claire Infante-Rivard, John Ioannidis, Astrid James, Giselle Jones, Bruno Ledergerber, Julian Little, Margaret May, David Moher, Hooman Momen, Alfredo Morabia, Hal Morgenstern, Cynthia D Mulrow, Fred Paccaud, Stuart J Pocock, Charles Poole, Martin Röösli, Dietrich Rothenbacher, Kenneth Rothman, Caroline Sabin, Willi Sauerbrei, Lale Say, James J Schlesselman, Jonathan Sterne, Holly Syddall, Jan P Vandenbroucke, Ian White, Susan Wieland, Hywel Williams, Guang Yong Zou.

Conflict of interest statement

We declare that we have no conflict of interest.

Acknowledgments

The workshop was funded by the European Science Foundation (ESF). Additional funding was received from the Medical Research Council Health Services Research Collaboration, and the National Health Services Research and Development Methodology Programme. We are grateful to Gerd Antes, Kay Dickersin, Shah Ebrahim, and Richard Lilford for supporting the STROBE initiative. We are grateful to the following institutions that have hosted working meetings of the coordinating group: Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland; Department of Social Medicine, University of Bristol, Bristol, UK; London School of Hygiene and Tropical Medicine, London, UK; Nordic Cochrane Centre, Copenhagen, Denmark; and Centre for Statistics in Medicine, Oxford, UK. We are grateful to six reviewers who provided helpful comments on a previous draft of this paper.

References

1 Glasziou P, Vandenbroucke JP, Chalmers I. Assessing the quality of research. *BMJ* 2004; **328:** 39–41.
2 Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996; **312:** 1215–18.
3 Papanikolaou PN, Christidi GD, Ioannidis JP. Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. *CMAJ* 2006; **174:** 635–41.
4 Jüni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ* 2001; **323:** 42–46.
5 Egger M, Schneider M, Davey Smith G. Spurious precision? Meta-analysis of observational studies. *BMJ* 1998; **316:** 140–44.
6 Pocock SJ, Collier TJ, Dandreo KJ, et al. Issues in the reporting of epidemiological studies: a survey of recent practice. *BMJ* 2004; **329:** 883.
7 Lee W, Bindman J, Ford T, et al. Bias in psychiatric case-control studies: literature survey. *Br J Psychiatry* 2007; **190:** 204–09.
8 Tooth L, Ware R, Bain C, Purdie DM, Dobson A. Quality of reporting of observational longitudinal research. *Am J Epidemiol* 2005; **161:** 280–88.
9 Bogardus ST Jr, Concato J, Feinstein AR. Clinical epidemiological quality in molecular genetic research: the need for methodological standards. *JAMA* 1999; **281:** 1919–26.
10 Anon. Guidelines for documentation of epidemiologic studies. Epidemiology Work Group of the Interagency Regulatory Liaison Group. *Am J Epidemiol* 1981; **114:** 609–13.
11 Rennie D. CONSORT revised – improving the reporting of randomized trials. *JAMA* 2001; **285:** 2006–07.
12 Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001; **357:** 1191–94.
13 Moher D, Altman DG, Schulz KF, Elbourne DR. Opportunities and challenges for improving the quality of reporting clinical research: CONSORT and beyond. *CMAJ* 2004; **171:** 349–50.
14 Plint AC, Moher D, Morrison A, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust* 2006; **185:** 263–67.
15 Egger M, Jüni P, Bartlett C. Value of flow diagrams in reports of randomized controlled trials. *JAMA* 2001; **285:** 1996–99.

16 Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999; **354:** 1896–900.

17 Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med* 2003; **138:** 40–44.

18 Vandenbroucke JP, von Elm E, Altman DG, et al, for the STROBE initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Med* 2007; **4:** e297.

19 Vandenbroucke JP, von Elm E, Altman DG, et al, for the STROBE initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Ann Intern Med* (in press).

20 Vandenbroucke JP von Elm E, Altman DG, et al, for the STROBE initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Epidemiology* (in press).

21 Ioannidis JP, Evans SJ, Gotzsche PC, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; **141:** 781–88.

22 Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004; **328:** 702–08.

23 Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA* 2006; **295:** 1152–60.

24 Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. *Ann Intern Med* 2006; **144:** 364–67.

25 Ioannidis JP, Gwinn M, Little J, et al. A road map for efficient and reliable human genome epidemiology. *Nat Genet* 2006; **38:** 3–5.

26 Ormerod AD. CONSORT your submissions: an update for authors. *Br J Dermatol* 2001; **145:** 378–79.

27 Schriger DL. Suggestions for improving the reporting of clinical research: the role of narrative. *Ann Emerg Med* 2005; **45:** 437–43.

28 Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol* 2007; **36:** 666–76.

29 Bartlett C, Sterne J, Egger M. What is newsworthy? Longitudinal study of the reporting of medical research in two British newspapers. *BMJ* 2002; **325:** 81–84.

Research Methods & Reporting

Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement

BMJ 2009; 339 doi: http://dx.doi.org/10.1136/bmj.b2535 (Published 21 July 2009) Cite this as: BMJ 2009;339:b2535

David Moher[1,2], Alessandro Liberati[3,4], Jennifer Tetzlaff[1], Douglas G Altman[5] for the PRISMA Group

[1] *Ottawa Methods Centre, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada*

[2] *Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada*

[3] *Università di Modena e Reggio Emilia, Modena, Italy*

[4] *Centro Cochrane Italiano, Istituto Ricerche Farmacologiche Mario Negri, Milan, Italy*

[5] *Centre for Statistics in Medicine, University of Oxford, Oxford, United Kingdom*

Correspondence to: dmoher@ohri.ca

- Accepted 5 June 2009

David Moher and colleagues introduce PRISMA, an update of the QUOROM guidelines for reporting systematic reviews and meta-analyses

Systematic reviews and meta-analyses have become increasingly important in health care. Clinicians read them to keep up to date with their specialty,**1 2** and they are often used as a starting point for developing clinical practice guidelines. Granting agencies may require a systematic review to ensure there is justification for further research,**3** and some medical journals are moving in this direction.**4** As with all research, the value of a systematic review depends on what was done, what was found, and the clarity of reporting. As with other publications, the reporting quality of systematic reviews varies, limiting readers' ability to assess the strengths and weaknesses of those reviews.

Several early studies evaluated the quality of review reports. In 1987 Mulrow examined 50 review articles published in four leading medical journals in 1985 and 1986 and found that none met all eight explicit scientific criteria, such as a quality assessment of included studies.**5** In 1987 Sacks and colleagues evaluated the adequacy of reporting of 83 meta-analyses on 23 characteristics in six domains.**6** Reporting was generally poor; between one and 14 characteristics were adequately reported (mean 7.7, standard deviation 2.7). A 1996 update of this study found little improvement.**7**

In 1996, to address the suboptimal reporting of meta-analyses, an international group developed a guidance called the QUOROM statement (QUality Of Reporting Of Meta-analyses), which focused on the reporting of meta-analyses of randomised controlled trials.**8** In this article, we summarise a revision of these guidelines, renamed PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses), which have been updated to address several conceptual and practical advances in the science of systematic reviews (see box).

Conceptual issues in the evolution from QUOROM to PRISMA

Completing a systematic review is an iterative process

The conduct of a systematic review depends heavily on the scope and quality of included studies: thus systematic reviewers may need to modify their original review protocol during its conduct. Any systematic review reporting guideline should recommend that such changes can be reported and explained without suggesting that they are inappropriate. The PRISMA statement (items 5, 11, 16, and 23) acknowledges this iterative process. Aside from Cochrane reviews, all of which should have a protocol, only about 10% of systematic reviewers report working from a protocol.**9** Without a protocol that is publicly accessible, it is difficult to judge between appropriate and inappropriate modifications.

Conduct and reporting of research are distinct concepts

This distinction is, however, less straightforward for systematic reviews than for assessments of the reporting of an individual study, because the reporting and conduct of systematic reviews are, by nature, closely intertwined. For example, the failure of a systematic review to report the assessment of the risk of bias in included studies may be seen as a marker of poor conduct, given the importance of this activity in the systematic review process.**10**

Study-level versus outcome-level assessment of risk of bias

For studies included in a systematic review, a thorough assessment of the risk of bias requires both a study-level assessment (such as adequacy of allocation concealment) and, for some features, a newer approach called outcome-level assessment. An outcome-level assessment involves evaluating the reliability and validity of the data for each important outcome by determining the methods used to assess them in each individual study.**11** The quality of evidence may differ across outcomes, even within a study, such as between a primary efficacy outcome, which is likely to be carefully and systematically measured, and the assessment of serious harms,**12** which may rely on spontaneous reports by investigators. This information should be reported to allow an explicit assessment of the extent to which an estimate of effect is correct.**11**

Importance of reporting biases

Different types of reporting biases may hamper the conduct and interpretation of systematic reviews. Selective reporting of complete studies (such as publication bias),**13** as well as the more recently empirically demonstrated "outcome reporting bias" within individual studies,**14 15** should be considered by authors when conducting a systematic review and reporting its results. Although the implications of these biases on the conduct and reporting of systematic reviews themselves are

unclear, some research has identified that selective outcome reporting may occur also in the context of systematic reviews.16

Terminology

The terminology used to describe a systematic review and meta-analysis has evolved over time. One reason for changing the name from QUOROM to PRISMA was the desire to encompass both systematic reviews and meta-analyses. We have adopted the definitions used by the Cochrane Collaboration.17 A systematic review is a review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to collect and analyse data from the studies that are included in the review. Statistical methods (meta-analysis) may or may not be used to analyse and summarise the results of the included studies. Meta-analysis refers to the use of statistical techniques in a systematic review to integrate the results of included studies.

Developing the PRISMA statement

A three-day meeting was held in Ottawa, Canada, in June 2005 with 29 participants, including review authors, methodologists, clinicians, medical editors, and a consumer. The objective of the Ottawa meeting was to revise and expand the QUOROM checklist and flow diagram as needed.

The executive committee completed the following tasks before the meeting: a systematic review of studies examining the quality of reporting of systematic reviews; a comprehensive literature search to identify methodological and other articles that might inform the meeting, especially in relation to modifying checklist items; and an international survey of review authors, consumers, and groups commissioning or using systematic reviews and meta-analyses (including the International Network of Agencies for Health Technology Assessment and the Guidelines International Network) to ascertain views of QUOROM, including the merits of the existing checklist items. The results of these activities were presented during the meeting and are summarised on the PRISMA website, www.prisma-statement.org/.

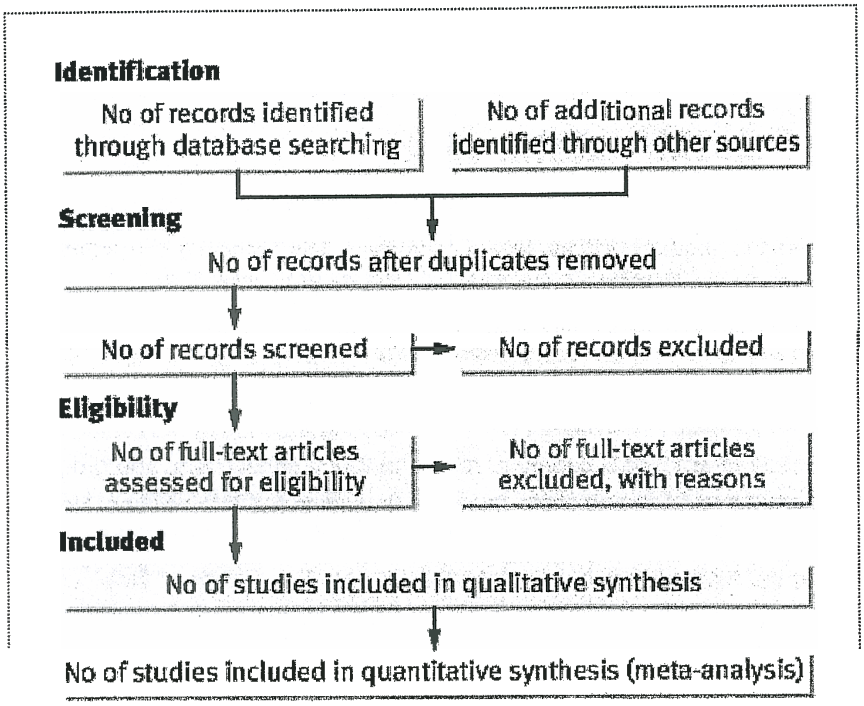
Only items deemed essential were retained or added to the checklist. Some additional items are nevertheless desirable, and review authors should include these, if relevant.18 For example, it is useful to indicate whether the systematic review is an update of a previous review19 and to describe any changes in procedures from those described in the original protocol.

Shortly after the meeting, a draft of the PRISMA checklist was circulated to the group, including those invited to the meeting but unable to attend. A disposition file was created containing comments and revisions from each respondent, and the checklist was subsequently revised 11 times. The group approved the checklist, flow diagram, and this summary paper.

Although no direct evidence was found to support retaining or adding some items, evidence from other domains was believed to be relevant. For example, item 5 asks authors to provide registration information about the systematic review, including a registration number if available. Although systematic review registration is not yet widely available,20 21 the participating journals of the International Committee of Medical Journal Editors22 now require all clinical trials to be registered in an effort to increase transparency and accountability.23 Those aspects are also likely to benefit systematic reviewers, possibly reducing the risk of an excessive number of reviews addressing the same question24 25 and providing greater transparency when updating systematic reviews.

The PRISMA statement

The PRISMA statement consists of a 27 item checklist (table 1⇓) and a four phase flow diagram (figure⇓) (also available as extra items on bmj.com for researchers to download and re-use). The aim of the PRISMA statement is to help authors improve the reporting of systematic reviews and meta-analyses. We have focused on randomised trials, but PRISMA can also be used as a basis for reporting systematic reviews of other types of research, particularly evaluations of interventions. PRISMA may also be useful for critical appraisal of published systematic reviews. However, the PRISMA checklist is not a quality assessment instrument to gauge the quality of a systematic review.



Flow of information through the different phases of a systematic review.

Table 1
Checklist of items to include when reporting a systematic review or meta-analysis

Section/topic	Item No	Checklist item	Reported on page No
Title			
Title	1	Identify the report as a systematic review, meta-analysis, or both	
Abstract			
Structured summary	2	Provide a structured summary including, as applicable, background, objectives, data sources, study eligibility criteria, participants, interventions, study appraisal and synthesis methods, results, limitations, conclusions and implications of key findings, systematic review registration number	

Section/topic	Item No	Checklist item	Reported on page No
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known	
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS)	
Methods			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (such as web address), and, if available, provide registration information including registration number	
Eligibility criteria	6	Specify study characteristics (such as PICOS, length of follow-up) and report characteristics (such as years considered, language, publication status) used as criteria for eligibility, giving rationale	
Information sources	7	Describe all information sources (such as databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched	
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated	
Study selection	9	State the process for selecting studies (that is, screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis)	
Data collection process	10	Describe method of data extraction from reports (such as piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators	
Data items	11	List and define all variables for which data were sought (such as PICOS, funding sources) and any assumptions and simplifications made	
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis	
Summary measures	13	State the principal summary measures (such as risk ratio, difference in means).	
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (such as I^2 statistic) for each meta-analysis	
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (such as publication bias, selective reporting within studies)	

Section/topic	Item No	Checklist item	Reported on page No
Additional analyses	16	Describe methods of additional analyses (such as sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified	
Results			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram	
Study characteristics	18	For each study, present characteristics for which data were extracted (such as study size, PICOS, follow-up period) and provide the citations	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome-level assessment (see item 12).	
Results of individual studies	20	For all outcomes considered (benefits or harms), present for each study (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals, ideally with a forest plot	
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see item 15)	
Additional analysis	23	Give results of additional analyses, if done (such as sensitivity or subgroup analyses, meta-regression) (see item 16)	
Discussion			
Summary of evidence	24	Summarise the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (such as health care providers, users, and policy makers)	
Limitations	25	Discuss limitations at study and outcome level (such as risk of bias), and at review level (such as incomplete retrieval of identified research, reporting bias)	
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research	
Funding			
Funding	27	Describe sources of funding for the systematic review and other support (such as supply of data) and role of funders for the systematic review	

Table 1
Checklist of items to include when reporting a systematic review or meta-analysis

From QUOROM to PRISMA

The new PRISMA checklist differs in several respects from the QUOROM checklist, and table 2⇓ lists the substantive specific changes. Generally, the PRISMA checklist "decouples" several items present in the QUOROM checklist and, where applicable, several checklist items are linked to improve consistency across the systematic review report.

Table 2
Substantive specific changes between the QUOROM checklist and the PRISMA checklist (a tick indicates the presence of the topic in QUOROM or PRISMA)

Section/topic and item	QUOROM	PRISMA	Comment
Abstract	√	√	QUOROM and PRISMA ask authors to report an abstract. However, PRISMA is not specific about format
Introduction:			
Objective		√	This new item (4) addresses the explicit question the review addresses using the PICO reporting system (which describes the participants, interventions, comparisons, and outcome(s) of the systematic review), together with the specification of the type of study design (PICOS); the item is linked to items 6, 11, and 18 of the checklist
Methods:			
Protocol		√	This new item (5) asks authors to report whether the review has a protocol and if so how it can be accessed
Search	√	√	Although reporting the search is present in both QUOROM and PRISMA checklists, PRISMA asks authors to provide a full description of at least one electronic search strategy (item 8). Without such information it is impossible to repeat the authors' search
Assessment of risk of bias in included studies	√	√	Renamed from "quality assessment" in QUOROM. This item (12) is linked to reporting this information in the results (item 19). The new concept of "outcome level" assessment has been introduced
Assessment of risk of bias across studies		√	This new item (15) asks authors to describe any assessments of risk of bias in the review, such as selective reporting within the included studies. This item is linked to reporting this information in the results (item 22)
Discussion	√	√	Although both QUOROM and PRISMA checklists address the discussion section, PRISMA devotes three items (24-26) to the discussion. In PRISMA the main types of limitations are explicitly stated and their discussion required
Funding		√	This new item (27) asks authors to provide information on any sources of funding for the systematic review.

Table 2
Substantive specific changes between the QUOROM checklist and the PRISMA checklist (a tick indicates the presence of the topic in QUOROM or PRISMA)

The flow diagram has also been modified. Before including studies and providing reasons for excluding others, the review team must first search the literature. This search results in records. Once these records have been screened and eligibility criteria applied, a smaller number of articles will remain. The number of included articles might be smaller (or larger) than the number of studies, because articles may report on multiple studies and results from a particular study may be published in several articles. To capture this information, the PRISMA flow diagram now requests information on these phases of the review process.

Endorsement

The PRISMA statement should replace the QUOROM statement for those journals that have endorsed QUOROM. We hope that other journals will support PRISMA; they can do so by registering on the PRISMA website. To emphasise to authors and others the importance of transparent reporting of systematic reviews, we encourage supporting journals to reference the PRISMA statement and include the PRISMA web address in their instructions to authors. We also invite editorial organisations to consider endorsing PRISMA and encourage authors to adhere to its principles.

The PRISMA explanation and elaboration paper

In addition to the PRISMA statement, a supporting explanation and elaboration document has been produced**26** following the style used for other reporting guidelines.**27 28 29** The process of completing this document included developing a large database of exemplars to highlight how best to report each checklist item, and identifying a comprehensive evidence base to support the inclusion of each checklist item. The explanation and elaboration document was completed after several face to face meetings and numerous iterations among several meeting participants, after which it was shared with the whole group for additional revisions and final approval. Finally, the group formed a dissemination subcommittee to help disseminate and implement PRISMA.

Discussion

The quality of reporting of systematic reviews is still not optimal.**9 30 31 32 33 34** In a recent review of 300 systematic reviews, few authors reported assessing possible publication bias,**9** even though there is overwhelming evidence for its existence**13** and its impact on the results of systematic reviews.**35** Even when the possibility of publication bias is assessed, there is no guarantee that systematic reviewers have assessed or interpreted it appropriately.**36** Although the absence of reporting such an assessment does not necessarily indicate that it was not done, reporting an assessment of possible publication bias is likely to be a marker of the thoroughness of the conduct of the systematic review.

Several approaches have been developed to conduct systematic reviews on a broader array of questions. For example, systematic reviews are now conducted to investigate cost effectiveness,**37** diagnostic**38** or prognostic questions,**39** genetic associations,**40** and policy making.**41** The general concepts and topics covered by PRISMA are relevant to any systematic review, not just those summarising the benefits and harms of a healthcare intervention. However, some modifications of the checklist items or flow diagram will be necessary in particular circumstances. For example, assessing the risk of bias is a key concept, but the items used to assess this in a diagnostic review are likely to focus on issues such as the spectrum of patients and the verification of disease status, which differ from reviews of interventions. The flow diagram will also need adjustments when reporting meta-analysis of individual patient data.**42**

We have developed an explanatory document to increase the usefulness of PRISMA.**26** For each checklist item, this document contains an example of good reporting, a rationale for its inclusion, and supporting evidence, including references, whenever possible. We believe this document will also serve as a useful resource for those teaching systematic review methodology. We encourage journals to include reference to the explanatory document in their instructions to authors.

Like any evidence based endeavour, PRISMA is a living document. To this end we invite readers to comment on the revised version, particularly the new checklist and flow diagram, through the PRISMA website. We will use such information to inform PRISMA's continued development.

Notes

Cite this as: *BMJ* 2009;338:b2535

Footnotes

- The following people contributed to the PRISMA statement: Doug Altman, Centre for Statistics in Medicine (Oxford); Gerd Antes, University Hospital Freiburg (Freiburg, Germany); David Atkins, Health Services Research and Development Service, Veterans Health Administration (Washington DC, USA); Virginia Barbour, *PLoS Medicine* (Cambridge, UK); Nick Barrowman, Children's Hospital of Eastern Ontario (Ottawa, Canada); Jesse A. Berlin, Johnson & Johnson Pharmaceutical Research and Development (Titusville NJ, USA); Jocalyn Clark, *PLoS Medicine* (at the time of writing, *BMJ*, London); Mike Clarke, UK Cochrane Centre (Oxford) and School of Nursing and Midwifery, Trinity College (Dublin, Ireland); Deborah Cook, Departments of Medicine, Clinical Epidemiology and Biostatistics, McMaster University (Hamilton, Canada); Roberto D'Amico, Università di Modena e Reggio Emilia (Modena, Italy) and Centro Cochrane Italiano, Istituto Ricerche Farmacologiche Mario Negri (Milan, Italy); Jonathan J Deeks, University of Birmingham (Birmingham); P J Devereaux, Departments of Medicine, Clinical Epidemiology and Biostatistics, McMaster University; Kay Dickersin, Johns Hopkins Bloomberg School of Public Health (Baltimore MD, USA); Matthias Egger, Department of Social and Preventive Medicine, University of Bern (Bern, Switzerland); Edzard Ernst, Peninsula Medical School (Exeter, UK); Peter C Gøtzsche, Nordic Cochrane Centre (Copenhagen, Denmark); Jeremy Grimshaw, Ottawa Hospital Research Institute (Ottawa, Canada); Gordon Guyatt, Departments of Medicine, Clinical Epidemiology and Biostatistics, McMaster University; Julian Higgins, MRC Biostatistics Unit (Cambridge, UK); John P A Ioannidis, University of Ioannina Campus (Ioannina, Greece); Jos Kleijnen, Kleijnen Systematic Reviews (York, UK) and School for Public Health and Primary Care (CAPHRI), University of Maastricht (Maastricht, Netherlands); Tom Lang, Tom Lang Communications and Training (Davis CA, USA); Alessandro Liberati, Università di Modena e Reggio Emilia, and Centro Cochrane Italiano, Istituto Ricerche Farmacologiche Mario Negri; Nicola Magrini, NHS Centre for the Evaluation of the Effectiveness of Health Care—CeVEAS (Modena, Italy); David McNamee, *Lancet* (London, UK); Lorenzo Moja, Centro Cochrane Italiano, Istituto Ricerche Farmacologiche Mario Negri; David Moher, Ottawa Methods Centre, Ottawa Hospital Research Institute (Ottawa, Canada); Cynthia Mulrow, *Annals of Internal Medicine* (Philadelphia PA, USA); Maryann Napoli, Center for Medical Consumers (New York, USA); Andy Oxman, Norwegian Health Services Research Centre (Oslo, Norway); Ba' Pham, Toronto Health Economics and Technology Assessment Collaborative (Toronto, Canada) (at the time of first meeting of the group, GlaxoSmithKline Canada,

Mississauga, Ontario); Drummond Rennie, University of California San Francisco (San Francisco CA, USA); Margaret Sampson, Children's Hospital of Eastern Ontario (Ottawa, Canada); Kenneth F Schulz, Family Health International (Durham NC, USA); Paul G Shekelle, Southern California Evidence Based Practice Center (Santa Monica CA, USA); Jennifer Tetzlaff, Ottawa Methods Centre, Ottawa Hospital Research Institute; David Tovey, *Cochrane Library*, Cochrane Collaboration (Oxford, UK) (at the time of first meeting of the group, *BMJ*, London); Peter Tugwell, Institute of Population Health, University of Ottawa (Ottawa, Canada).

- Author contributions: ICMJE criteria for authorship read and met—DM. Agree with the recommendations—DM, AL, JT, DGA. Wrote the first draft of the paper—DM, AL, DGA. Contributed to the writing of the paper—DM, AL, JT, DGA. Participated in regular conference calls, identified the participants, secured funds, planned the meeting, participated in the meeting, and drafted the manuscript—DM, AL, DGA. Participated in identifying the evidence base for PRISMA, refining the checklist, and drafting the manuscript—JT.

- Funding: PRISMA was funded by the Canadian Institutes of Health Research; Università di Modena e Reggio Emilia, Italy; Cancer Research UK; Clinical Evidence BMJ Knowledge; the Cochrane Collaboration; and GlaxoSmithKline, Canada. AL is funded, in part, through grants of the Italian Ministry of University (COFIN-PRIN 2002 prot 2002061749 and COFIN-PRIN 2006 prot 2006062298). DGA is funded by Cancer Research UK. DM is funded by a University of Ottawa Research Chair. None of the sponsors had any involvement in the planning, execution, or writing of the PRISMA documents. No funder played a role in drafting this manuscript.

- Competing interests: None declared.

- Provenance and peer review: Not commissioned; externally peer reviewed.

- In order to encourage dissemination of the PRISMA statement, this article is freely accessible on bmj.com and will also be published in *PLoS Medicine*, *Annals of Internal Medicine*, *Journal of Clinical Epidemiology*, and *Open Medicine*. The authors jointly hold the copyright of this article. For details on further use, see the PRISMA website (www.prisma-statement.org/).

References

1. Oxman AD, Cook DJ, Guyatt GH. Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. *JAMA*1994;**272**:1367-71.
2. Swingler GH, Volmink J, Ioannidis JP. Number of published systematic reviews and global burden of disease: database analysis. *BMJ*2003;**327**:1083-4.
3. Canadian Institutes of Health Research. Randomized controlled trials registration/application checklist (12/2006). 2006. www.cihr-irsc.gc.ca/e/documents/rct_reg_e.pdf (accessed 19 May 2009).
4. Young C, Horton R. Putting clinical trials into context. *Lancet*2005;**366**:107.
5. Mulrow CD. The medical review article: state of the science. *Ann Intern Med*1987;**106**:485-8.
6. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analysis of randomized controlled trials. *N Engl J Med*1987;**316**:450-5.

7. Sacks HS, Reitman D, Pagano D, Kupelnick B. Meta-analysis: an update. *Mt Sinai J Med*1996;**63**:216-24.
8. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF, for the QUOROM group. Improving the quality of reporting of meta-analysis of randomized controlled trials: The QUOROM statement. *Lancet*1999;**354**:1896-1900.
9. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. *PLoS Med*2007;**4**:e78, doi:10.1371/journal.pmed.0040078.
10. Moja LP, Telaro E, D'Amico R, Moschetti I, Coe L, Liberati A. Assessment of methodological quality of primary studies by systematic reviews: results of the metaquality cross sectional study. *BMJ*2005;**330**:1053-5.
11. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al, for the GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*2008;**336**:924-6.
12. Schunemann HJ, Jaeschke R, Cook DJ, Bria WF, El-Solh AA, et al, for the ATS Documents Development and Implementation Committee. An official ATS statement: grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. *Am J Respir Crit Care Med*2006;**174**:605-14.
13. Dickersin K. Publication bias: recognizing the problem, understanding its origins and scope, and preventing harm. In: Rothstein HR, Sutton AJ, Borenstein M, eds. *Publication bias in meta-analysis—prevention, assessment and adjustments*. Chichester: John Wiley, 2005:11-33.
14. Chan AW, Hrobjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA*2004;**291**:2457-65.
15. Chan AW, Krleza-Jeric K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ*2004;**171**:735-40.
16. Silagy CA, Middleton P, Hopewell S. Publishing protocols of systematic reviews: comparing what was done to what was planned. *JAMA*2002;**287**:2831-4.
17. Green S, Higgins J, eds. Glossary. *Cochrane Handbook for Systematic Reviews of Interventions 4.2.5* [updated May 2005]. www.cochrane.org/resources/glossary.htm (accessed 19 May 2009).
18. Strech D, Tilburt J. Value judgments in the analysis and synthesis of evidence. *J Clin Epidemiol*2008;**61**:521-4.
19. Moher D, Tsertsvadze A. Systematic reviews: when is an update an update? *Lancet*2006;**367**:881-3.
20. University of York. Centre for Reviews and Dissemination, 2009. www.york.ac.uk/inst/crd/ (accessed 19 May 2009).
21. Joanna Briggs Institute. Protocols & work in progress, 2008. www.joannabriggs.edu.au/pubs/systematic_reviews_prot.php (accessed 19 May 2009).
22. De Angelis C, Drazan JM, Frizelle FA, Haug C, Hoey J, et al, for the International Committee Medical Journal Editors. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *CMAJ*2004;**171**:606-7.
23. Whittington CJ, Kendall T, Fonagy P, Cottrell D, Cotgrove A, et al. Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet*2004;**363**:1341-5.
24. Bagshaw SM, McAlister FA, Manns BJ, Ghali WA. Acetylcysteine in the prevention of contrast-induced nephropathy: a case study of the pitfalls in the evolution of evidence. *Arch Intern Med*2006;**166**:161-6.
25. Biondi-Zoccai GG, Lotrionte M, Abbate A, Testa L, Remigi E, et al. Compliance with QUOROM and quality of reporting of overlapping meta-analyses on the role of acetylcysteine in the prevention of contrast associated nephropathy: case study. *BMJ*2006;**332**:202-9.

26. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al, for the PRISMA Group. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*2009;**339**:b2700.

27. Altman DG, Schulz KR, Moher D, Egger M, Davidoff F, et al, for the CONSORT group. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*2001;**134**:663-94.

28. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, et al, for the STARD group. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD explanation and elaboration. *Ann Intern Med*2003;**138**:W1-12.

29. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, et al, for the STROBE initiative. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Ann Intern Med*2007;**147**:W163-94.

30. Bhandari M, Morrow F, Kulkarni AV, Tornetta P. Meta-analyses in orthopaedic surgery: a systematic review of their methodologies. *J Bone Joint Surg Am*2001;83-A:15-24.

31. Kelly KD, Travers A, Dorgan M, Slater L, Rowe BH. Evaluating the quality of systematic reviews in the emergency medicine literature. *Ann Emerg Med*2001;**38**:518-26.

32. Richards D. The quality of systematic reviews in dentistry. *Evid Based Dent*2004;**5**:17.

33. Choi PT, Halpern SH, Malik N, Jadad AR, Tramer MR, et al. Examining the evidence in anesthesia literature: a critical appraisal of systematic reviews. *Anesth Analg*2001;**92**:700-9.

34. Delaney A, Bagshaw SM, Ferland A, Manns B, Laupland KB. A systematic evaluation of the quality of meta-analyses in the critical care literature. *Crit Care*2005;**9**:R575-82.

35. Sutton AJ. Evidence concerning the consequences of publication and related biases. In: Rothstein HR, Sutton AJ, Borenstein M, eds. *Publication bias in meta-analysis—prevention, assessment and adjustments*. Chichester: John Wiley, 2005:175-92.

36. Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. *BMJ*2006;**333**:597-600.

37. Ladabaum U, Chopra CL, Huang G, Scheiman JM, Chernew ME, et al. Aspirin as an adjunct to screening for prevention of sporadic colorectal cancer: a cost-effectiveness analysis. *Ann Intern Med*2001;**135**:769-81.

38. Deeks JJ. Systematic reviews in health care: systematic reviews of evaluations of diagnostic and screening tests. *BMJ*2001;**323**:157-62.

39. Altman DG. Systematic reviews of evaluations of prognostic variables. *BMJ*2001;**323**:224-8.

40. Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet*2001;**29**:306-9.

41. Lavis J, Davies H, Oxman A, Denis J, Golden-Biddle K, et al. Towards systematic reviews that inform health care management and policy-making. *J Health Serv Res Policy*2005;**10**:35-48.

42. Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data. Cochrane Working Group. *Stat Med*1995;**14**:2057-79.

Basic Statistical Reporting for Articles Published in Biomedical Journals: The "Statistical Analyses and Methods in the Published Literature" or The SAMPL Guidelines"

Thomas A. Lang[a] and **Douglas G. Altman**[b]
[a] Principal, Tom Lang Communications and Training International
[b] Director, Centre for Statistics in Medicine, Oxford University

Have they reflected that the sciences founded on observation can only be promoted by statistics? . . . If medicine had not neglected this instrument, this means of progress, it would possess a greater number of positive truths, and stand less liable to the accusation of being a science of unfixed principles, vague and conjectural.

Jean-Etienne Dominique Esquirol, an early French psychiatrist,
quoted in The Lancet, 1838 [1]

Introduction

The first major study of the quality of statistical reporting in the biomedical literature was published in 1966 [2]. Since then, dozens of similar studies have been published, every one of which has found that large proportions of articles contain errors in the application, analysis, interpretation, or reporting of statistics or in the design or conduct of research. (See, for example, references 3 through 19.) Further, large proportions of these errors are serious enough to call the authors' conclusions into question [5,18,19]. The problem is made worse by the fact that most of these studies are of the world's leading peer-reviewed general medical and specialty journals.

Although errors have been found in more complex statistical procedures [20,21,22], paradoxically, many errors are in basic, not advanced, statistical methods [23]. Perhaps advanced methods are suggested by consulting statisticians, who then competently perform the analyses, but it is also true that authors are far more likely to use only elementary statistical methods, if they use any at all [23-26]. Still, articles with even major errors continue to pass editorial and peer review and to be published in leading journals.

The truth is that the problem of poor statistical reporting is long-standing, widespread, potentially serious, concerns mostly basic statistics, and yet is largely unsuspected by most readers of the biomedical literature [27].

More than 30 years ago, O'Fallon and colleagues recommended that "Standards governing the content and format of statistical aspects should be developed to guide authors in the preparation of manuscripts" [28]. Despite the fact that this call has since been echoed by several others (17,18,29-32), most journals have still not included in their Instructions for Authors more than a paragraph or two about reporting statistical methods [33]. However, given that many statistical errors concern basic statistics, a

Lang T, Altman D. Basic statistical reporting for articles published in clinical medical journals: the SAMPL Guidelines. In: Smart P, Maisonneuve H, Polderman A (eds). *Science Editors' Handbook*, European Association of Science Editors, 2013.

comprehensive—and comprehensible—set of reporting guidelines might improve how statistical analyses are documented.

In light of the above, we present here a set of statistical reporting guidelines suitable for medical journals to include in their Instructions for Authors. These guidelines tell authors, journal editors, and reviewers how to report basic statistical methods and results. Although these guidelines are limited to the most common statistical analyses, they are nevertheless sufficient to prevent most of the reporting deficiencies routinely found in scientific articles; they may also help to prevent some reporting errors by focusing attention on key points in the analyses.

Unlike many of other guidelines, the SAMPL guidelines were not developed by a formal consensus-building process, but they do draw considerably from published guidelines [27,34-37]. In addition, a comprehensive review of the literature on statistical reporting errors reveals near universal agreement on how to report the most common methods [27].

Statistical analyses are closely related to the design and activities of the research itself. However, our guidelines do not address the issues related to the design and conduct of research. Instead, we refer readers to the EQUATOR Network website (www.equator-network.org) where guidelines for reporting specific research designs can be found. (For example, see the CONSORT [38], TREND [39], STROBE [40]) These guidelines for reporting methodologies all include items on reporting statistics, but the guidelines presented here are more specific and complement, not duplicate, those in the methodology guidelines.

We welcome feedback and anticipate the need to update this guidance in due course.

Reporting Basic Statistical Analyses and Methods in the Published Literature: The SAMPL Guidelines for Biomedical Journals

Guiding Principles for Reporting Statistical Methods and Results

Our first guiding principle for statistical reporting comes from The International Committee of Medical Journal Editors, whose Uniform Requirements for Manuscripts Submitted to Biomedical Journals include the following excellent statement about reporting statistical analyses:

> "**Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results.** [Emphasis added.] When possible, quantify findings and present them with appropriate indicators of measurement error or uncertainty (such as confidence intervals). Avoid relying solely on statistical hypothesis testing, such as *P* values, which fail to convey important information about effect size. References for the design of the study and statistical methods should be to standard works when possible (with pages stated). Define statistical terms, abbreviations, and most symbols. Specify the computer software used" [33,41].

Our second guiding principle for statistical reporting is to **provide enough detail that the results can be incorporated into other analyses.** In general, this principle requires reporting the descriptive statistics from which other statistics are derived, such as the numerators and denominators of percentages, especially in risk, odds, and hazards ratios. Likewise, P values are not sufficient for re-analysis. Needed instead are descriptive statistics for the variables being compared, including sample size of the groups involved, the estimate (or "effect size") associated with the P value, and a measure of precision for the estimate, usually a 95% confidence interval.

General Principles for Reporting Statistical Methods

Preliminary analyses

- Identify any statistical procedures used to modify raw data before analysis. Examples include mathematically transforming continuous measurements to make distributions closer to the normal distribution, creating ratios or other derived variables, and collapsing continuous data into categorical data or combining categories.

Primary analyses

- Describe the purpose of the analysis.
- Identify the variables used in the analysis and summarize each with descriptive statistics.
- When possible, identify the smallest difference considered to be clinically important.
- Describe fully the main methods for analyzing the primary objectives of the study.
- Make clear which method was used for each analysis, rather than just listing in one place all the statistical methods used.
- Verify that that data conformed to the assumptions of the test used to analyze them. In particular, specify that 1) skewed data were analyzed with non-parametric tests, 2) paired data were analyzed with paired tests, and 3) the underlying relationship analyzed with linear regression models was linear.
- Indicate whether and how any allowance or adjustments were made for multiple comparisons

(performing multiple hypothesis tests on the same data).

- If relevant, report how any outlying data were treated in the analysis.
- Say whether tests were one- or two-tailed and justify the use of one-tailed tests.
- Report the alpha level (e.g., 0.05) that defines statistical significance.
- Name the statistical package or program used in the analysis.

Supplementary analyses

- Describe methods used for any ancillary analyses, such as sensitivity analyses, imputation of missing values, or testing of assumptions underlying methods of analysis.
- Identify post-hoc analyses, including unplanned subgroup analyses, as exploratory.

General Principles for Reporting Statistical Results

Reporting numbers and descriptive statistics

- Report numbers—especially measurements—with an appropriate degree of precision. For ease of comprehension and simplicity, round to a reasonable extent. For example, mean age can often be rounded to the nearest year without compromising either the clinical or the statistical analysis. If the smallest meaningful difference on a scale is 5 points, scores can be reported as whole numbers; decimals are not necessary.
- Report total sample and group sizes for each analysis.
- Report numerators and denominators for all percentages.
- Summarize data that are approximately normally distributed with means and standard deviations (SD). Use the form: mean (SD), not mean ± SD.
- Summarize data that are not normally distributed with medians and interpercentile ranges, ranges, or both. Report the upper and lower boundaries of interpercentile ranges and the minimum and maximum values of ranges, not just the size of the range.
- Do NOT use the standard error of the mean (SE) to indicate the variability of a data set. Use standard deviations, inter-percentile ranges, or ranges instead. (The SE is an inferential statistic—it is about a 68% confidence interval—not a descriptive statistic.)
- Display data in tables or figures. Tables present exact values, and figures provide an overall assessment of the data.[42,43]

Reporting risk, rates, and ratios

- Identify the type of rate (e.g., incidence rates; survival rates), ratio (e.g., odds ratios; hazards ratios), or risk (e.g., absolute risks; relative risk differences), being reported.
- Identify the quantities represented in the numerator and denominator (e.g., the number of men with prostate cancer divided by the number of men in whom prostate cancer can occur).
- Identify the time period over with each rate applies.
- Identify any unit of population (that is, the unit multiplier: e.g., x 100; x 10,000) associated with the rate.
- Consider reporting a measure of precision (a confidence interval) for estimated risks, rates, and ratios.

Reporting hypothesis tests

- State the hypothesis being tested.
- Identify the variables in the analysis and summarize the data for each variable with the appropriate descriptive statistics.
- If possible, identify the minimum difference considered to be clinically important.
- For equivalence and non-inferiority studies, report the largest difference between groups that will still be accepted as indicating biological equivalence (the equivalence margin).
- Identify the name of the test used in the analysis. Report whether the test was one- or two-tailed (justify the use of one-tailed tests) and for paired or independent samples.
- Confirm that the assumptions of the test were met by the data.
- Report the alpha level (e.g., 0.05) that defines statistical significance.
- At least for primary outcomes, such as differences or agreement between groups, diagnostic sensitivity, and slopes of regression lines, report a measure of precision, such as the 95% confidence interval.
- Do NOT use the standard error of the mean (SE) to indicate the precision of an estimate. The SE is essentially a 68% confidence coefficient: use the 95% confidence coefficient instead.
- Although not preferred to confidence intervals, if desired, *P* values should be reported as equalities when possible and to one or two decimal places (e.g., $P = 0.03$ or 0.22 not as inequalities: e.g., $P < 0.05$). Do NOT report "NS"; give the actual P value. The smallest *P* value that need be reported is $P < 0.001$, save in studies of genetic associations.
- Report whether and how any adjustments were made for multiple statistical comparisons.
- Name the statistical software package used in the analysis.

Reporting association analyses

- Describe the association of interest.
- Identify the variables used and summarize each with descriptive statistics.
- Identify the test of association used.
- Indicate whether the test was one- or two-tailed. Justify the use of one-tailed tests.
- For *tests* of association (e.g., a *chi*-square test), report the *P* value of the test (because association is defined as a statistically significant result).
- For *measures* of association (i.e., the *phi* coefficient), report the value of the coefficient and a confidence interval. Do not describe the association as low, moderate, or high unless the ranges for these categories have been defined. Even then, consider the wisdom of using these categories given their biological implications or realities.
- For primary comparisons, consider including the full contingency table for the analysis.
- Name the statistical package or program used in the analysis.

Reporting correlation analyses

- Describe the purpose of the analysis.
- Summarize each variable with the appropriate descriptive statistics.
- Identify the correlation coefficient used in the analysis (e.g., Pearson, Spearman).
- Confirm that the assumptions of the analysis were met.

Lang T, Altman D. Statistical Analyses and Methods in the Published Literature: the SAMPL Guidelines.

- Report the alpha level (e.g., 0.05) that indicates whether the correlation coefficient is statistically significant.
- Report the value of the correlation coefficient. Do not describe correlation as low, moderate, or high unless the ranges for these categories have been defined. Even then, consider the wisdom of using these categories given their biological implications or realities.
- For primary comparisons, report the (95%) confidence interval for the correlation coefficient, whether or not it is statistically significant.
- For primary comparisons, consider reporting the results as a scatter plot. The sample size, correlation coefficient (with its confidence interval), and *P* value can be included in the data field.
- Name the statistical package or program used in the analysis.

Reporting regression analyses

- Describe the purpose of the analysis.
- Identify the variables used in the analysis and summarize each with descriptive statistics.
- Confirm that the assumptions of the analysis were met. For example, in linear regression indicate whether an analysis of residuals confirmed the assumptions of linearity.
- If relevant, report how any outlying values were treated in the analysis.
- Report how any missing data were treated in the analyses.
- For either simple or multiple (multivariable) regression analyses, report the regression equation.
- For multiple regression analyses: 1) report the alpha level used in the univariate analysis; 2) report whether the variables were assessed for a) colinearity and b) interaction; and 3) describe the variable selection process by which the final model was developed (e.g., forward-stepwise; best subset).
- Report the regression coefficients (beta weights) of each explanatory variable and the associated confidence intervals and *P* values, preferably in a table.
- Provide a measure of the model's "goodness-of-fit" to the data (the coefficient of determination, r^2, for simple regression and the coefficient of multiple determination, R^2, for multiple regression).
- Specify whether and how the model was validated.
- For primary comparisons analyzed with simple linear regression analysis, consider reporting the results graphically, in a scatter plot showing the regression line and its confidence bounds. Do not extend the regression line (or the interpretation of the analysis) beyond the minimum and maximum values of the data.
- Name the statistical package or program used in the analysis.

Reporting analyses of variance (ANOVA) or of covariance (ANCOVA)

- Describe the purpose of the analysis.
- Identify the variables used in the analysis and summarize each with descriptive statistics.
- Confirm that the assumptions of the analysis were met. For example, indicate whether an analysis of residuals confirmed the assumptions of linearity.
- If relevant, report how any outlying data were treated in the analysis.
- Report how any missing data were treated in the analyses.
- Specify whether the explanatory variables were tested for interaction, and if so how these interactions were treated.
- If appropriate, in a table, report the *P* value for each explanatory variable, the test statistics and, where applicable, the degrees of freedom for the analysis.

- Provide an assessment of the goodness-of-fit of the model to the data, such as R^2.
- Specify whether and how the model was validated.
- Name the statistical package or program used in the analysis.

Reporting survival (time-to-event) analyses

- Describe the purpose of the analysis.
- Identify the dates or events that mark the beginning and the end of the time period analyzed.
- Specify the circumstances under which data were censored.
- Specify the statistical methods used to estimate the survival rate.
- Confirm that the assumptions of survival analysis were met.
- For each group, give the estimated survival probability at appropriate follow-up times, with confidence intervals, and the number of participants at risk for death at each time. It is often more helpful to plot the cumulative probability of not surviving, especially when events are not common.
- Reporting median survival times, with confidence intervals, is often useful to allow the results to be compared with those of other studies.
- Consider presenting the full results in a graph (e.g., a Kaplan-Meier plot) or table.
- Specify the statistical methods used to compare two or more survival curves.
- When comparing two or more survival curves with hypothesis tests, report the *P* value of the comparison
- Report the regression model used to assess the associations between the explanatory variables and survival or time-to-event.
- Report a measure of risk (e.g., a hazard ratio) for each explanatory variable, with a confidence interval.

Reporting Bayesian analyses

- Specify the pre-trial probabilities ("priors").
- Explain how the priors were selected.
- Describe the statistical model used.
- Describe the techniques used in the analysis.
- Identify the statistical software program used in the analysis.
- Summarize the posterior distribution with a measure of central tendency and a credibility interval
- Assess the sensitivity of the analysis to different priors.

References

1. Esquirol JED. Cited in: Pearl R. Introduction to Medical Biometry and Statistics. Philadelphia: WB Saunders, 1941.
2. Schor S, Karten I. Statistical evaluation of medical journal manuscripts. JAMA. 1966;195:1123-8.
3. Nagele P. Misuse of standard error of the mean (SEM) when reporting variability of a sample. A critical evaluation of four anaesthesia journals. Brit J Anaesth. 2003; 90: 514-6.
4. Neville JA, Lang W, Fleischer AB Jr. Errors in the Archives of Dermatology and the Journal of the American Academy of Dermatology from January through December 2003. *Arch Dermatol*. 2006; 142: 737-40.
5. Glantz SA. Biostatistics: how to detect, correct and prevent errors in the medical literature. Circulation. 1980;61:1-7.
6. Lionel ND, Herxheimer A. Assessing reports of therapeutic trials. BMJ. 1970;3:637- 40.

7. Altman DG. Statistics in medical journals: developments in the 1980s. Stat Med. 1991;10:1897-913.
8. White S J. Statistical errors in papers in the British Journal of Psychiatry. Br J Psychiatr. 1979;135:336-42.
9. Gore SM, Jones IG, Rytter EC. Misuse of statistical methods: critical assessment of articles in BMJ from January to March 1976. BMJ. 1977;1:85-7.
10. Scales CD Jr, Norris RD, Peterson BL, et al. Clinical research and statistical methods in the urology literature. J Urol. 2005;174:1374-19.
11. Kurichi JE, Sonnad SS. Statistical methods in the surgical literature. J Am Col Surg. 2006;202:476-84.
12. Gardner MJ, Altman DG, Jones DR, Machin D. Is the statistical assessment of papers submitted to the British Medical Journal effective? BMJ. 1983;286:1485-8.
13. Bakker M, Wicherts JM. The (mis)reporting of statistical results in psychology journals. *Behav Res.* 2011;43:666-78.
14. Avram MJ, Shanks CA, Dykes MH, et al. Statistical methods in anesthesia articles: an evaluation of two American journals during two six-month periods. Anesth Analg. 1985;64:607-11.
15. Godfrey K. Comparing the means of several groups. N Engl J Med.1985;313:1450-6.
16. A survey of three medical journals. N Engl J Med. 1987;317:426-32.
17. Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. N Engl J Med. 1987 Aug 13;317(7):426-32.
18. Murray GD. Statistical aspects of research methodology. Br J Surg. 1991;78:777-81.
19. Yancy JM. Ten rules for reading clinical research reports [Editorial]. Am J Surg. 1990;159:553-9.
20. Burton A, Altman DG. Missing covariate data within cancer prognostic studies: a review of current reporting and proposed guidelines. *Br J Cancer* 2004;91:4-8.
21. Mackinnon A. The use and reporting of multiple imputation in medical research – a review. *J Intern Med* 2010;268:586-93.
22. Schwarzer G, Vach W, Schumacher M. On the misuses of artificial neural networks for prognostic and diagnostic classification in oncology. *Stat Med* 2000;19:541-61.
23 George SL. Statistics in medical journals: a survey of current policies and proposals for editors. Med Pediatr Oncol. 1985;13:109-12.
24. Emerson JD, Colditz GA. The statistical content of published medical research: some implications for biomedical education. Med Educ. 1985:19(3);248-255. DOI: 10.1111/j.1365-2923.1985.tb01315.x
25. Golden J, Zhu W, Sayre JW. A review of the statistical analysis used in papers published in Clinical Radiology and British Journal of Radiology. Clin Radiol. 1996;51(1):47- 50.
26. Lee CM, Soin HK, Einarson TR. Statistics in the Pharmacy Literature. Ann Pharmacother. 2004; 38(9):1412-1418. DOI 10.1345/aph.1D493
27. Lang T, Secic M. How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers, Second edition. Philadelphia: American College of Physicians, 2006.
28. O'Fallon JR, Duby SD, Salsburg DS, et al. Should there be statistical guidelines for medical research papers? Biometrics, 1978;34:687-95.
29. Shott S. Statistics in veterinary research. J Am Vet Med Assoc. 1985;187:138-41.
30. Hayden GF. Biostatistical trends in Pediatrics: implications for the future. Pediatrics. 1983;72:84-7.
31. Altman DG, Bland JM. Improving doctors' understanding of statistics. J R Statist Soc A. 1991;154:223-67.
32. Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. BMJ. 1983; 286:1489-93.
33. Bailar JC 3rd, Mosteller F. Guidelines for statistical reporting in articles for medical journals. Amplifications and explanations. Ann Intern Med. 1988 108(2):266-73.
34. Bond GR, Mintz J, McHugo GJ. Statistical guidelines for the Archives of PM&R. *Arch Phys Med Rehabil* 1995;76:784-7.
35. Wilkinson L and the Task Force on Statistical Inference. Statistical methods in psychology journals. Guidelines and explanations. *Am Psychologist* 1999;54:594-604.
36. Curran-Everett D, Benos DJ; American Physiological Society. Guidelines for reporting statistics in journals published by the American Physiological Society. *Am J Physiol Endocrinol Metab* 2004;287:E189-91. (plus other journals)
37. Curran-Everett D, Benos DJ. Guidelines for reporting statistics in journals published by the American Physiological Society: the sequel. *Adv Physiol Educ* 2007;31:295-8.
38. Moher D, Schulz K, Altman DG, for the CONSORT Group. CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. Ann Intern Med. 2001;134:657-62.
39. Des Jarlais DC, Lyles C, Crepaz N, Trend Group. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Public Health. 2004; 94(3):361-6. PMID: 14998794
40. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. Ann Intern Med. 2007; 147(8):573-577. PMID: 17938396
41. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals: writing and editing for biomedical publication, 2011. www.icmje.org. Accessed December 12, 2012.
42. Schriger DL, Arora S, Altman DG. The content of medical journal instructions for authors. *Ann Emerg Med* 2006;48:743-749, 749.e1-4.
43. Lang T. How to Write, Publish, and Present in the Health Sciences: A Guide for Clinicians and Laboratory Researchers. Philadelphia: American College of Physicians, 2010.

Volume 3, number 2

What is critical appraisal?

Sponsored by an educational grant from AVENTIS Pharma

Alison Hill BSC FFPHM FRCP Director, and **Claire Spittlehouse** BSc Business Manager, Critical Appraisal Skills Programme, Institute of Health Sciences, Oxford

- **Critical appraisal** is the process of **systematically** examining research evidence to assess its **validity, results** and **relevance** before using it to inform a decision.
- Critical appraisal is an essential part of **evidence-based clinical practice** that includes the process of systematically **finding, appraising** and **acting** on evidence of effectiveness.
- Critical appraisal allows us to make sense of research evidence and thus begins to close the gap between research and practice.
- **Randomised controlled trials** can minimise bias and use the most appropriate design for studying the effectiveness of a specific intervention or treatment.
- **Systematic reviews** are particularly useful because they usually contain an explicit statement of the objectives, materials and methods, and should be conducted according to explicit and reproducible methodology.
- Randomised controlled trials and systematic reviews are not automatically of good quality and should be appraised critically.

Prescribing information is on page 8

What is critical appraisal?

Critical appraisal is one step in the process of evidence-based clinical practice. Evidence-based clinical practice is 'an approach to decision making in which the clinician uses the best evidence available, in consultation with the patient, to decide the option which suits the patient best'.[1] To determine what is the 'best' evidence, we need critical appraisal skills that will help us to understand the methods and results of research and to assess the quality of the research. Most research is not perfect, and critical appraisal is not an exact science – it will not give us the 'right' answer. But it can help us to decide whether we think a reported piece of research is good enough to be used in decision making.

There are many factors that come into play when making healthcare decisions – research evidence is just one of them. If research has flaws, it is up to readers to use their critical appraisal skills to decide whether this affects the usefulness of the paper in influencing their decision.

Pros of critical appraisal in practice

- Critical appraisal provides a systematic way of assessing the validity, results and usefulness of published research papers.
- Together with skills in finding research evidence and changing practice as a result of research, critical appraisal is the route to closing the gap between research and practice[1] and as such makes an essential contribution to improving healthcare quality.
- Critical appraisal encourages objective assessment of the usefulness of information – critical appraisal skills are applied to published research, but all evidence should be appraised to weigh up its usefulness.
- Critical appraisal skills are not difficult to develop. Critical appraisal is a common sense approach to reading, and user-friendly tools are available to help anyone develop these skills.

Cons of critical appraisal in practice

- Critical appraisal can be time-consuming initially, although with time it becomes the automatic way to look at research papers.
- Critical appraisal does not always provide the reader with the 'easy' answer or the answer one might have hoped for; it may highlight that a favoured intervention is in fact ineffective.
- Critical appraisal can be dispiriting if it highlights a lack of good evidence – it may take determination to persist with an area of interest when access to good research in the area is limited.

Appraising randomised controlled trials

Box 1[3] (opposite) provides a checklist of questions for critically appraising ***randomised controlled trials (RCTs).*** The RCT is the most appropriate research design for studying the effectiveness of a specific intervention or treatment.[2] In an RCT, participants are randomly assigned to two (or more) groups: one (or more) experimental group(s) receiving the intervention that is being tested, and a comparison or control group receiving a placebo or an alternative treatment. The two (or more) groups are then followed up to see what differences result. Randomisation ensures that the groups differ only in the intervention given, so any difference between the outcomes in each group can be attributed to the intervention.

RCTs' methodology *can* minimise accidental or intentional bias, but this does not automatically mean that every RCT is of good quality. We must *critically appraise* individual studies to assess the validity of their methods. Once we are happy that the methods were sound, then we can look at what the results tell us and consider whether we can apply them to our own population.

This method of critically appraising an RCT can be applied to the paper about the Heart Outcomes Prevention Evaluation

Box 1. 12 questions to help you make sense of a trial. Adapted from Guyatt *et al*[3]

A. Are the results of the study valid?

Screening questions

1. *Did the trial address a clearly focused research question?*
   Tip: a research question should be 'focused' in terms of:
   - The population studied
   - The intervention given
   - The outcomes considered.
2. *Did the authors use the right type of study?*
   Tip: the right type of study would:
   - Address the research question
   - Have an appropriate study design.

Is it worth continuing?

Detailed questions

3. *Was the assignment of patients to treatments randomised?*
   Tip: consider if this was done appropriately.
4. *Were all of the patients who entered the trial properly accounted for at its conclusion?*
   Tip: look for:
   - The completion of follow-up
   - Whether patients were analysed in the groups to which they were randomised.
5. *Were patients, health workers and study personnel 'blind' to treatment?*
   Tip: this is not always possible, but consider if it was possible – was every effort made to ensure 'blinding'?
6. *Were the groups similar at the start of the study?*
   Tip: think about other factors that might effect the outcome such as age, sex, social class.
7. *Aside from the experimental intervention, were the groups treated equally?*
   Tip: for example, were they reviewed at the same time intervals.

B. What are the results?

8. *How large was the treatment effect?*
9. *How precise was the estimate of the treatment effect?*
   Tip: look for the confidence limits.

C. Will the results help locally?

10. Can the results be applied to the local population?
    Tip: consider whether the patients covered by the trial are likely to be very different from your population.
11. Were all clinically important outcomes considered?
12. Are the benefits worth the harms and costs?

(HOPE) study,[4] which investigates the effects of ramipril on cardiovascular events in high-risk patients. A brief critical appraisal of the methods of the HOPE paper (addressed by questions 1–7) is included to demonstrate how critical appraisal tools can be used.

How should these questions be used?

Questions one and two are screening questions, and any research paper should give the reader the answers to these two questions on the first page, ideally in the abstract. If it is not possible to answer 'yes' to both these

questions, we would normally not continue to read the paper as it is unlikely to be helpful. All questions in sections A and C should be answered with either 'yes', 'no' or 'can't tell', which should be justified with evidence from the text; section B requires text answers.

Section A: are the results of the study valid?

The first seven questions address the methods used in the trial. If the methods are robust, we can then look at the results of the trial and decide if the results can be useful for us. If the research methods are flawed, then this may invalidate the trial.

Screening questions:

1. Did the trial address a clearly focused research question?

It is important to be clear about the focus of the question – does the research set out to answer a question that will be useful to you and can help to inform your decision? Unfocused research cannot provide you with reliable answers. If it is not focused it is not worth spending time appraising it.

2. Did the authors use the right type of study?

Having determined the research question in question 1 above, you now need to assess if the paper uses an appropriate type of research method that is likely to answer that question. If the research question is 'does intervention x produce the desired measurable outcome in a specific population?', then the best type of single research study to address this is an RCT. Alternatively, if the research question were 'what are the views, beliefs and perceptions of a specific population?', then the best type of research study to address this would be a qualitative study.

For example, the HOPE paper[4] asks the question: *does ramipril [the intervention] prevent myocardial infarction, stroke or death from cardiovascular causes [the outcomes] in patients who were at high risk for cardiovascular events but who did not have left ventricular dysfunction or heart disease [the population]?* This question is best addressed by an RCT.

Is it worth continuing?

If we have established that the paper has a clear research question that can be addressed by an RCT, we need to continue through this checklist of questions and consider the quality of the methods used.

Detailed questions:

3. Was the assignment of patients to treatments randomised?

The randomisation method should be robust to reduce the possibility of bias within the allocation of participants to treatment and control groups. The most robust types of randomisation are by computer-generated numbers or tables of random numbers: the participants have a 50:50 chance of being in either group, and both the participants and the clinicians are completely unaware before randomisation which group the participants will be in. Quasi randomisation is less robust and often involves randomising by alternate participants, day of the week or alternate months.

In the HOPE paper we are told that the study participants were randomly assigned to groups, but we are not given any details of the randomisation method, which would have given us more confidence in the methods. A previous paper is referenced for details of the study design.[5] (This previous paper tells us that patients were randomised internationally by a telephone call to a central office and assigned to groups in a 2x2 factorial design.)

4. Were all of the patients who entered the trial properly accounted for at its conclusion?

It is important to account for the participants in a trial who were **lost to follow up** – that is, those who withdrew from the trial before its conclusion. Those who are lost are not necessarily representative of the whole group, so their loss may distort the results. It is also important to consider how the participants were analysed – was ***intention-to-treat analysis*** used? That is, analysing the participants in the groups to which they were randomised – this is important because if participants move between groups, the random generation of groups will have been affected and, again, this may distort the results.

At the end of the HOPE study, 99.9% of randomised patients were accounted for – this is a very good level of follow-up. We must also bear in mind that 32% of patients had discontinued treatment at some stage during

the trial and 28% permanently discontinued. There was also a run-in phase before the RCT began, which resulted in the exclusion of 10% of patients before randomisation due to non-compliance, side-effects, abnormal serum creatinine or potassium levels, or withdrawal of consent. The implications of this would need to be addressed in **section C**, which addresses the generalisability or external validity of the study.

5. Were patients, health workers and study personnel 'blind' to treatment?

If the people involved in the study are **blind** (that is, they are not aware of who is in the treatment or control groups), this reduces the possibility of bias. To blind patients and health workers, the intervention and placebo must appear to be identical, which is not always possible. However, the study personnel can always be blinded; they do not need to be aware of whether the patients they assess or data that they are working with is from the treatment or control group.

In the HOPE study, we could assume that the patients were blinded as we are told that they were given the drug or a 'matching placebo'. It would have been straightforward to have blinded the health workers administering the drug and the study personnel who analysed the data as well, but this is not discussed in this paper. (In the previous paper[5] it is stated that 'emergency unblinding is available centrally and locally but will only be done when absolutely necessary and after a check list is completed by telephone call to the project office', which implies that there was blinding.)

6. Were the groups similar at the start of the study?

If the sample is large enough and the randomisation method robust, then the make-up of the treatment and control groups should be very similar and the only difference should be the intervention or control or alternative treatment. This is important if we are to be reasonably sure that the outcome of the trial is due to the intervention and not any other factors.

In the HOPE study, the randomisation process appears to have been successful, as the baseline characteristics of the patients in the two groups are very similar. This is clearly displayed in a table.

7. Aside from the experimental intervention, were the groups treated equally?

Again, it is important to be confident that the groups were treated equally, so that we can attribute the outcome of the trial to the intervention. It is possible for groups to be treated differently if the health workers or study personnel are not blinded.

In the HOPE study, we know that the groups were followed up over the same time period and we have no reason to believe that the groups were *not* treated equally. If we assume that everyone involved in the trial was blinded, then the groups must have been treated equally.

Section B: what are the results?

Questions 8 and 9 address the results presented in the paper.

8. How large was the treatment effect?

The outcome measures used differ between papers. Often a measure of ***relative risk*** is used. The chances of a particular outcome being observed in an individual is called the risk (risk can refer to something good or bad). By comparing the risk in the experimental and control groups, a measure of relative risk can be calculated.[6] A relative risk of 1 occurs when the incidences are the same in the two groups.[7] If we hope that the intervention would lead to more of the outcome measured (for example, reduction of symptoms) then we want a relative risk of more than 1; if we hope that the intervention would lead to less of the outcome measured (for example, myocardial infarction, stroke or death from cardiovascular causes, which were the primary outcomes considered in the HOPE trial) then we want a relative risk of less than 1. Results may also be presented as NNTs (numbers needed to treat).[8]

9. How precise was the estimate of the treatment effect?

There will always be some doubt about the result (or best estimate), as a trial only looks at a sample of the population. The **confidence interval (CI)** indicates the range of doubt around the best estimate.[6] Results are often also presented with **p values.** These describe the probability that a particular result has happened by chance.[6] If the p value is less

than 0.05, then this is usually described as being **statistically significant** and means that the results are unlikely to be due to chance.

Section C: will the results help locally?

Any weaknesses in the methods or in the presentation of the results should be borne in mind when considering how useful the results are locally.

10. Can the results be applied to the local population?

It is important to consider if there are any differences between the participants in the trial and the local population that would make it impossible to apply the results locally. Participants dropping out of the study must be considered as their loss may distort the results. The patients lost from the HOPE study are discussed in question 4.

11. Were all clinically important outcomes considered?

A single trial cannot address all the important outcomes that we are interested in, but consider if the paper has answered the original research question and whether any other important outcomes have been highlighted or missed out.

12. Are the benefits worth the harms and costs?

Financial information is not normally given in a trial, but we need to consider what the negative effects could be and whether these are outweighed by the benefits. Other research, such as an economic evaluation, might help with the cost implications.

Healthcare decisions are not usually made purely on the basis of one trial. There are many other factors that influence decisions, and there may be many trials that together can provide more conclusive evidence. Systematic reviews can provide valuable information.

Appraising systematic reviews

Box 2[9] (opposite) provides a checklist for appraising systematic reviews. Reviews collect together primary research and summarise their results and conclusions. Systematic reviews are particularly useful because they usually contain an explicit statement of objectives, materials and methods, and should have been conducted according to explicit and reproducible methodology. But, as with RCTs, systematic reviews should be critically appraised by users so they can decide for themselves whether their methods are valid, assess what the results are saying and decide whether these results can be applied locally. Further information on systematic reviews can be found in 'What is a systematic review?'.[10]

Appraising other types of studies

In addition to RCTs and systematic reviews, it is increasingly being recognised that other types of studies contribute to evidence-based decision-making. These include cohort and case-control studies, economic evaluations, studies on diagnostic tests and qualitative studies. Checklists for these types of studies are available from the Critical Appraisal Skills Programme (CASP).[2,6,11] The JAMA user guides and the Centre for Evidence-based Medicine are other important sources of appraisal checklists.[12,13]

References

1. Gray JAM. *Evidence-based healthcare: how to make health policy and management decisions*. Edinburgh: Churchill Livingstone, 1997.
2. Critical Appraisal Skills Programme and HCLU. *Evidence-based health care: an open learning resource for healthcare professionals*. Oxford: Update Software, 1999.
3. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II: how to use an article about therapy or prevention. *JAMA* 1993; **270:** 2598–2601 and **271:** 59–63.
4. The Heart Outcomes Prevention Evaluation (HOPE) Study Investigators. Effects of an angiotensin-converting-enzyme inhibitor, ramipril, on cardiovascular events in high-risk patients. *N Engl J Med* 2000; **342:** 145–153.
5. The HOPE Study Investigators. The HOPE (Heart Outcomes Prevention Evaluation) study: the design of a large, simple randomized trial of an angiotensin-converting enzyme inhibitor (ramipril) and vitamin E in patients at high-risk of cardiovascular events. *Can J Cardiol* 1996; **12**(2): 127–137.
6. Critical Appraisal Skills Programme and HCLU. *Evidence-based health care. A computer aided learning resource.* Oxford: Update Software, 1999.
7. Kirkwood BR. Essentials of medical statistics. Oxford: Blackwell Science, 1998.
8. Moore A, McQuay HJ. *What is an NNT?* London: Hayward Medical Communications, 1997.
9. Oxman AD *et al.* Users' guides to the medical literature. VI: how to use an overview. *JAMA* 1994; **272**(17): 1367–1371.
10. Davies HT, Crombie IK. *What is a systematic review?* London: Hayward Medical Communications, 1998.
11. CASP. http://www.casp.org.uk
12. Centre for Evidence-based Medicine. http://cebm.jr2.ox.ac.uk/
13. Canadian Centers for Health Evidence. http://www.cche.net/

Box 2. Ten questions to help you make sense of a systematic review. Adapted from Oxman *et al* [9]

Three broad issues need to be considered when appraising research:

A Are the results of the study valid?
B What are the results?
C Will the results help locally?

The questions below are designed to help you think about these issues systematically

- The first two questions are screening questions and can be answered quickly. If the answer to both is 'yes', it is worth continuing.
- There is a fair degree of overlap between several of the questions.
- You are asked to record a 'yes', 'no' or 'can't tell' to most of the questions.
- A number of tips are given after each question. These are designed to remind you why the question is important.

A. Are the results of the review valid?

Screening questions

1. *Did the review address a clearly focused research question?*
   Tip: a research question should be 'focused' in terms of:
   - The population studied
   - The intervention given or exposure
   - The outcomes considered.
2. *Did the review include the right type of studies?*
   Tip: these would:
   - Address the review's research question
   - Have an appropriate study design.

Is it worth continuing?

Detailed questions

3. *Did the reviewers try to identify all relevant studies?*
   Tip: look for:
   - Which bibliographic databases were used
   - Follow-up from reference lists
   - Personal contact with experts
   - Search for unpublished studies
   - Search for non-English language studies.
4. *Did the reviewers assess the quality of the included studies?*
   Tip: a clear predetermined strategy should be used to determine which studies are included. Look for:
   - A scoring system
   - More that one assessor.
5. *If the results of the studies have been combined, was it reasonable to do so?*
   Tip: consider whether:
   - The results of each study are clearly displayed
   - The results were similar from study to study (look for tests of heterogeneity)
   - The reasons for any variations in results are discussed.
   Tip: think about other factors that might effect the outcome such as age, sex, social class.

B. What are the results?

6. *What are the main results of the review?*
   Tip: consider:
   - How the results are expressed (for example, odds ratio, relative risk and so on)
   - What the results are.
7. *Could the results be due to chance?*
   Tip: look for tests of statistical significance (p values) and confidence intervals (CIs).

C. Will the results help locally?

8. *Can the results be applied to the local population?*
   Tip: consider whether:
   - The population sample covered by the review could be sufficiently different from your population to cause concern
   - Your local setting is likely to differ much from that of the review.
9. *Were all important outcomes considered?*
   Tip: consider outcomes from the point of view of the:
   - Individual
   - Policy makers and practitioners
   - Family/carers
   - Wider community.
10. *Should policy of practice change as a result of teh evidence contained in this review?*
    Tip: consider whether the benefits are worth the harms and costs.

Tritace prescribing information

Presentation: Capsules containing 1.25mg, 2.5mg, 5mg or 10mg ramipril. **Indications:** Reducing the risk of myocardial infarction, stroke, cardiovascular death or need for revascularisation procedures in patients of 55 years or more who have clinical evidence of cardiovascular disease (previous MI, unstable angina or multivessel CABG or multivessel PTCA), stroke or peripheral vascular disease. Reducing the risk of myocardial infarction, stroke, cardiovascular death or need for revascularisation procedures in diabetic patients of 55 years or more who have one or more of the following clinical findings: hypertension (systolic blood pressure >160mmHg or diastolic blood pressure >90mmHg); high total cholesterol (>5.2mmol/l); low HDL (< 0.9mmol/l); current smoker; known microalbuminuria; clinical evidence of previous vascular disease. Mild to moderate hypertension. Congestive heart failure as adjunctive therapy to diuretics with or without cardiac glycosides. Reduction in mortality in patients surviving acute MI with clinical evidence of heart failure. **Dosage and administration:** Reduction in risk of MI, stroke, cardiovascular death or need for revascularisation procedure: The initial dose is 2.5mg Tritace o.d.. Depending on tolerability, the dose should be gradually increased. It is recommended that this dose is doubled after about 1 week of treatment then, after a further 3 weeks, increased to 10mg. The usual maintenance dose is 10mg Tritace o.d.. Patients stabilised on lower doses of Tritace for other indications where possible should be titrated to 10mg Tritace o.d.. Hypertension: Initial dose 1.25mg titrated up to a maximum of 10mg o.d. according to response. Usual dose 2.5mg or 5mg o.d.. Stop diuretic therapy 2 - 3 days before starting Tritace and resume later if required. Congestive heart failure: Initial dose 1.25mg o.d. titrated up to a maximum of 10mg daily according to response. Doses above 2.5mg daily can be given o.d. or b.d.. Post Myocardial Infarction: Initiate treatment between day 3 and day 10 following AMI. Initially 2.5mg b.d. increasing to 5mg b.d. after 2 days. Assessment of renal function is recommended prior to initiation. Reduced maintenance dose may be required in impaired renal function. Monitor patients with impaired liver function. In the elderly the dose should be titrated according to need. Not recommended for children. **Contraindications:** Hypersensitivity to ramipril or excipients, history of angioneurotic oedema, haemodynamically relevant renal artery stenosis, hypotensive or haemodynamically unstable patients, pregnancy, lactation. **Precautions:** Do not use in aortic or mitral valve stenosis or outflow obstruction. Assess renal function before and during use, as there is a risk of impairment of renal function. Use with caution during surgery or anaesthesia. Hyperkalaemia. Do not use in patients using polyacrylonitrile (AN69) dialysis membranes or during low-density lipoprotein apheresis with dextran sulphate. Agranulocytosis and bone marrow depression seen rarely with ACE inhibitors as well as a reduction in red cell count, haemoglobin and platelet content. Symptomatic hypotension may occur after initial dose or increase in dose, especially in salt/volume depleted patients. **Drug interactions:** Combination with diuretics, NSAIDs, adrenergic blocking drugs or other antihypertensive agents may potentiate antihypertensive effect. Risk of hyperkalaemia when used with agents increasing serum potassium. May enhance the effect of antidiabetic agents. May increase serum lithium concentrations. **Side effects:** Dizziness, headache, weakness, disturbed balance, nervousness, restlessness, tremor, sleep disorders, confusion, loss of appetite, depressed mood, anxiety, paraesthesiae, taste changes, muscle cramps & joint pain, erectile impotence, reduced sexual desire, fatigue, cough, hypersensitivity reactions; pruritus, rash, shortness of breath, fever, cutaneous and mucosal reactions, Raynauds phenomenon, gastrointestinal disturbances, jaundice, hepatitis, impaired renal function, angioneurotic oedema, pancreatitis, eosinophilia and vasculitis. Symptomatic hypotension, myocardial infarction or cerebrovascular accident possibly secondary to severe hypotension in high-risk patients, chest pain, palpitations, rhythm disturbances, angina pectoris may occur. Use with caution and closely monitor patients with impaired liver function. Reduced serum sodium levels, elevated blood urea nitrogen and serum creatinine. Pre-existing proteinuria may deteriorate.
Basic NHS cost: 28 x 1.25mg capsules £5.30; 28 x 2.5mg capsules £7.51; 28 x 5mg capsules £9.55; 28 x 10mg capsules £13.00
Marketing authorisation numbers: 1.25mg 13402/0021, 2.5mg 13402/0022, 5mg 13402/0023, 10mg 13402/0024
Legal category: POM
Marketing authorisation holder: Aventis Pharma, 50 Kings Hill Avenue, Kings Hill, West Malling, Kent, ME19 4AH
Date of preparation: July 2000

This publication, along with the others in the series, is available on the internet at www.evidence-based-medicine.co.uk

Sponsored by an educational grant from *AVENTIS Pharma*

Published by Hayward Medical Communications, a division of Hayward Group plc.

TRI1320101
Date of preparation: February 2001

12 questions to help you make sense of cohort study

How to use this appraisal tool

Three broad issues need to be considered when appraising a cohort study:

- **Are the results of the study valid?** **(Section A)**
- **What are the results?** **(Section B)**
- **Will the results help locally?** **(Section C)**

The 12 questions on the following pages are designed to help you think about these issues systematically. The first two questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions.

There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

There will not be time in the small groups to answer them all in detail!

(A) Are the results of the study valid?

Screening Questions

1. Did the study address a clearly focused issue? ❒ Yes ❒ Can't tell ❒ No

HINT: A question can be 'focused' In terms of

- The population studied
- The risk factors studied
- The outcomes considered
- Is it clear whether the study tried to detect a beneficial or harmful effect?

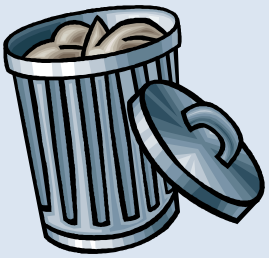
---

2. Was the cohort recruited in an acceptable way? ❒ Yes ❒ Can't tell ❒ No

HINT: Look for selection bias which might compromise the generalisibility of the findings:

- Was the cohort representative of a defined population?
- Was there something special about the cohort?
- Was everybody included who should have been included?

Is it worth continuing?



Detailed questions

3. Was the exposure accurately measured to minimise bias?

❒ Yes ❒ Can't tell ❒ No

HINT: Look for measurement or classification bias:

- Did they use subjective or objective measurements?
- Do the measurements truly reflect what you want them to (have they been validated)?
- Were all the subjects classified into exposure groups using the same procedure

---

4. Was the outcome accurately measured to minimise bias?

❒ Yes ❒ Can't tell ❒ No

HINT: Look for measurement or classification bias:

- Did they use subjective or objective measurements?
- Do the measures truly reflect what you want them to (have they been validated)?
- Has a reliable system been established for detecting all the cases (for measuring disease occurrence)?
- Were the measurement methods similar in the different groups?
- Were the subjects and/or the outcome assessor blinded to exposure (does this matter)?

5. (a) Have the authors identified all important confounding factors? ❑ Yes ❑ Can't tell ❑ No

List the ones you think might be important, that the author missed.

(b) Have they taken account of the confounding factors in the design and/or analysis? ❑ Yes ❑ Can't tell ❑ No

List:

HINT: Look for restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

---

6. (a) Was the follow up of subjects complete enough? ❑ Yes ❑ Can't tell ❑ No

(b) Was the follow up of subjects long enough? ❑ Yes ❑ Can't tell ❑ No

HINT: Consider

- The good or bad effects should have had long enough to reveal themselves
- The persons that are lost to follow-up may have different outcomes than those available for assessment
- In an open or dynamic cohort, was there anything special about the outcome of the people leaving, or the exposure of the people entering the cohort?

(B) What are the results?

7. What are the results of this study?

HINT: Consider

- What are the bottom line results?
- Have they reported the rate or the proportion between the exposed/unexposed, the ratio/the rate difference?
- How strong is the association between exposure and outcome (RR,)?
- What is the absolute risk reduction (ARR)?

---

8. How precise are the results?

HINT: Look for the range of the confidence intervals, if given.

---

9. Do you believe the results?

❒ Yes ❒ Can't tell ❒ No

HINT: Consider

- Big effect is hard to ignore!
- Can it be due to bias, chance or confounding?
- Are the design and methods of this study sufficiently flawed to make the results unreliable?
- Bradford Hills criteria (e.g. time sequence, dose-response gradient, biological plausibility, consistency)

(C) Will the results help locally?

10. Can the results be applied to the local population? ❑ Yes ❑ Can't tell ❑ No

HINT: Consider whether

- A cohort study was the appropriate method to answer this question
- The subjects covered in this study could be sufficiently different from your population to cause concern
- Your local setting is likely to differ much from that of the study
- You can quantify the local benefits and harms

---

11. Do the results of this study fit with other available evidence? ❑ Yes ❑ Can't tell ❑ No

---

12. What are the implications of this study for practice?

HINT: Consider

- One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making
- For certain questions observational studies provide the only evidence
- Recommendations from observational studies are always stronger when supported by other evidence

GUIDELINES FOR COMPLETING A RESEARCH PROTOCOL FOR OBSERVATIONAL STUDIES

The aim of this guide is to help researchers write a research study protocol for an observational study. The guide will take you through each section of the protocol giving advice and examples of the information required in that section. This is a guide only and for those requiring more information on particular topics; some useful references are given at the end of this document. We also recommend the St. George's on line handbook (now hosted at York University) at http://www-users.york.ac.uk/~mb55/guide/describ.htm and there are some links to that document in this guide.

Many of the methodological aspects of designing a research study and writing the protocol can benefit from the advice of a statistician. Such advice should be sought at an early stage and is available for UCL/UCLH/RFH researchers through the Biostatistics group at the Joint UCLH/UCL/RFH Biomedical Research Unit.

1. Title Page

1.1 Title

It is useful to specify both a full title and short title

- The full title should include summary study design, exposure and outcome of interest, patient population and setting.
- The short title is a summary of this
- The titles specified must be consistent across all documents relevant to the study.

1.2 Names (titles), roles and contact details of:

- Authors, investigators, experts and advisors involved in the study
- Sponsor & monitor – as agreed with Chief / principal investigator's employer and the host Trust
- Study site(s), clinical laboratory(s), technical departments and institutions involved in the study

1.3 Protocol details

- Version number
- final / draft
- Date

2. Signature Page

Signatures of all healthcare professionals involved in the study

3. Contents Page

4. List of Abbreviations

All abbreviations used should be listed and defined

5. Summary

In this section the problem being addressed should be introduced to "set the scene".

- Aim(s) and reason for the study
- A brief description of the study groups, design as well as the exposure / outcomes of interest.
- Primary and secondary objectives

- Brief description of methods

6. Background

The background should include a **clear explanation of the main research question** along with a full literature review, a detailed justification for the study and discussion of its feasibility.

6.1 Literature Search and Review

Precise details of the literature search completed should be given here.

- The nature of any unpublished work should also be documented here.
- Help with this aspect your intended research can often be sought form a clinical librarian. A good literature search at the beginning of the research process is invaluable and will steer the researcher in the most appropriate direction.
- Your review should make reference to relevant papers, unpublished works as well as clinical experience.

6.2 Justification

- A statement indicating the size of the problem (and effect on the health service) and why the study is appropriate
- Explain what the potential benefits are – to patients and the health service
- Explain what your study will add to the body of evidence already available.
- Discuss the feasibility of the study in terms of subject and data availability as well as length.

7. Specific aim(s) of the study

- Describe the primary research question
- Clearly defined objectives in terms of measurable endpoints
- Distinguish primary and secondary objectives

8. Study Design

The two most commonly used designs for observational studies are (A) case-control studies (including nested case-control studies) and (B) cohort studies. In the former, the study groups are chosen on the basis of their disease or outcome of interest. In a cohort study the comparison groups are identified according to an exposure of interest. A description of the design should be given, along with details of any matching and blinding used. Cross-sectional studies are also discussed in (C).