Reporting Guidelines Based on Principles of Evidence-Based Medicine

International Association of Veterinary Editors
1 June 2010    Geneva, Switzerland

Erik von Elm, MD MSc FMH
German Cochrane Center, Univ. of Freiburg / Germany
Swiss Paraplegic Research, Nottwil / Switzerland
Seminar Health Sciences & Policy, Univ. of Lucerne / Switzerland
vonelm@cochrane.de
Can you trust the research published in medical journals?

“Mostly I believe you can. But not always … some of the most newsworthy reports published by our leading journals spontaneously combust under the gentlest of scrutiny.”

Guardian 16 Sept 2004

Richard Horton
Editor-in-Chief The Lancet
Biomedical research & publication

- Biomedical research should advance scientific knowledge and – directly or indirectly – lead to improvements in prevention or management of illness.

- Publications are usually the only tangible evidence that a study was done, how it was done, and what the findings were.
Knowledge generation cycle

Research

Translation

Dissemination

Publication
Purpose of a research publication

- **Researchers** might read it to help plan a similar study or as part of a systematic review
  - They need a clear understanding of what was done exactly

- **Practioners** might read it to learn how to care for their patients better
  - They need a clear understanding of what was found

- **Policy makers** might read it to make decision about services & specific interventions
  - They need a clear understanding of what can be concluded
Knowledge generation cycle

Gap between what is done and what is reported
Common problems with reporting

- Non-reporting / underreporting of whole studies
- Delayed reporting
- Omission of crucial information in description of research methods (e.g. study interventions)
- Presenting data (e.g. in graphs) in confusing or misleading ways
  - particularly important for presenting benefits and harms
- Inadequate statistical reporting
- Selective reporting of some outcomes (but not others)
- Omissions or mis-interpretation of results in abstracts or full articles
Underreporting of research

Bluemle et al. (J Med Ethics 2008, 34:e20)

- 299 study protocols submitted to REC Freiburg i.Brsg. in year 2000
- Electronic searches for publications & survey of applicants in March – July 2007
- Publication outcome of 225 completed studies (75%)
  - 109 / 225 (48%) published
- Subgroup of randomised trials (46%): 
  - 54 / 103 (52%) published
Reporting of study methods (1): interventions

Glasziou et al. (BMJ 2008)

- Assessed descriptions of treatments in 80 published articles: 55 randomised trials & 25 systematic reviews published in *Evidence-Based Medicine*

- In 41 articles essential elements of interventions were missing

- Only 3 / 25 systematic reviews provided intervention description sufficient for implementation
Reporting of study methods (2): trial methodology

Chan & Altman (Lancet 2005)

- 519 randomised trials published in Dec 2000 & indexed in PubMed

**Failure to report key aspects of trial conduct:**
- 73% Sample size calculation
- 55% Defined primary outcome(s)
- 60% Whether blinded or not
- 79% Method of random sequence generation
- 82% Method of allocation concealment
Reporting of results (1): selective reporting

Accumulating empirical evidence of two major threats to the medical literature:

- **Study publication bias** – studies with less interesting findings are less likely to be published

- **Outcome reporting bias** – results included within published reports are selected to favour those with statistically significant results
Reporting of results (2): study publication bias

Example of antidepressants trials (Turner NEJM 2008)

74 FDA-registered trials (12564 part.) on 12 drugs were assessed
23 / 74 trials (3449 part.) remained unpublished

38 viewed as ‘positive’ by FDA:
  – 37 published
  – 1 unpublished

36 viewed as ‘negative’ by FDA:
  – only 3 published as ‘negative’
  – 11 published giving impression of ‘positive’ results
  – 22 unpublished

Look at published literature: 94% ‘positive’ trials,
at FDA reports: 51% ‘positive’ trials.
Reporting of results (3): outcome reporting bias

Protocol

Primary outcome: % of participants with Score X<3 at 1 year

Publication

Primary outcome: % of participants with Score X<3 at 1 year

\[ P \geq 0.05 \]
Reporting of results (3): outcome reporting bias

Protocol

Primary outcome:
% of participants with Score X<3 at 1 year

Publication

Primary outcome:
% of participants dead or dependent at 1 year

P<0.05
Reporting of results (4): outcome reporting bias

- **Comparison of protocols and publications**
  - 102 RCTs submitted to a Danish Ethics committee in 1994-95
    - Protocols and subsequent journal articles (122 articles)
    - Questionnaire sent to all authors

- **Frequent discrepancies between the protocol and the published report**
  - Specification of primary outcomes: 51/82 (62%)

[Chan et al, *JAMA* 2004]
Reporting of other study types

- Problems with reporting have been studied most frequently for randomised trials

- Similar concerns apply to all types of biomedical research:
  - Diagnostic accuracy studies
  - Observational studies (e.g. case-control studies, cohort studies)
  - Prognostic studies
  - Qualitative studies
  - Systematic reviews
  - etc.
Impact of poor reporting (1)

Poor reporting is a serious problem e.g. in systematic reviews of available evidence:

“The biggest problem was the quality of reporting, which did not allow us to judge the important methodological items ...”

“Data reporting was poor. 15 trials met the inclusion criteria for this review but only 4 could be included as data were impossible to use in the other 11.”

(Quotations from Cochrane Reviews)
Impact of poor reporting (2)

- "Users" of literature cannot assess reliability of individual studies
  - Methods may not be adequately described
  - Methodological weaknesses may not be apparent

- They cannot assess a body of evidence

- Consequences for
  - other researchers
  - health professionals in clinical and community context
  - people living with a health condition
Whose fault is it ?!

Poor reporting might be **collective problem of different groups involved in publication process:**

- **Researchers (authors)** may not know what information to include in a research article
- **Editors** may not know what information should be included
- **Peer reviewers** may not know how to evaluate a manuscript
A dialogue between an editor and Socrates about peer review
Peer review: some questions from Socrates…

*Socrates:*

“Let me try to sum up what you have told me so far. Judging the quality of manuscripts submitted to scientific journals is a difficult and demanding task. Editors often entrust this task to people who haven’t had any training in how to do it and (...) they may not be very good at it. Further, they may well be rivals or, on the other hand, friends or protégés of the authors so it is hard for them to be impartial.

The task takes them several hours – time that they might prefer to spend doing something else. They are rarely acknowledged or paid anything for their trouble. However bad they say the paper is in their report, the authors will probably get it into print somewhere.”

What does it tell us about peer review?

- Peer review is difficult and only partly successful
- Reviewers (and editorial staff) are often unable to eliminate errors
- Readers cannot assume that all papers published in peer reviewed journals are scientifically sound

⇒ Transparent reporting is critical to identify problems in submitted manuscripts & published articles
What can be done?

Guidance on research methods: GCP, GEP...

General guidance on scientific writing & Author Instructions
What can be done?

- Research
- Guidance on research methods: GCP, GEP...
- Translation
- Dissemination
- Publication
- Reporting guidelines
- Guidance on scientific writing & Author Instructions
An early call for guidance on reporting

- “This leads one to consider if it is possible, in planning a trial, in reporting the results, or in assessing the published reports of trials, to apply criteria which must be satisfied ...”

- “A basic principle can be set up that ... it is at least as important to describe the techniques employed and the conditions in which the experiment was conducted, as to give the detailed statistical analysis of results.”

Key aspects of reporting guidelines

- Established by international collaborative groups incl. researchers, methodologists and editors

- Specify a minimum set of items recommended for a clear and transparent account of what was done and what was found in a research study

- Emphasis on issues that might introduce bias into biomedical research

- Checklist of items based on evidence if available. If not, consensus opinion of experts.
**Principles of evidence-based medicine**

**Levels of evidence by Sackett et al. (modified version):**

<table>
<thead>
<tr>
<th>Level</th>
<th>Research Design</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>Randomized controlled trial (RCT)</td>
<td>Randomized controlled trial, PEDro score ≥ 6. Includes within subjects comparison with randomized conditions and crossover designs</td>
</tr>
<tr>
<td>Level 2</td>
<td>RCT</td>
<td>Randomized controlled trial, PEDro score &lt; 6.</td>
</tr>
<tr>
<td>Level 2</td>
<td>Prospective controlled trial</td>
<td>Prospective controlled trial (not randomized)</td>
</tr>
<tr>
<td>Level 2</td>
<td>Cohort</td>
<td>Prospective longitudinal study using at least 2 similar groups with one exposed to a particular condition.</td>
</tr>
<tr>
<td>Level 3</td>
<td>Case control</td>
<td>A retrospective study comparing conditions, including historical controls</td>
</tr>
<tr>
<td>Level 3</td>
<td>Pre-post</td>
<td>A prospective trial with a baseline measure, intervention, and a post-test using a single group of subjects.</td>
</tr>
<tr>
<td>Level 4</td>
<td>Post-test</td>
<td>A prospective post-test with two or more groups – intervention, then post-test (no pre-test or baseline measurement) using a single group of subjects.</td>
</tr>
<tr>
<td>Level 4</td>
<td>Case Series</td>
<td>A retrospective study usually collecting variables from a chart review.</td>
</tr>
<tr>
<td>Level 5</td>
<td>Observational</td>
<td>Study using cross-sectional analysis to interpret relations.</td>
</tr>
<tr>
<td>Level 5</td>
<td>Clinical Consensus</td>
<td>Expert opinion without explicit critical appraisal, or based on physiology, biomechanics or &quot;first principles&quot;</td>
</tr>
<tr>
<td>Level 5</td>
<td>Case Report</td>
<td>Pre-post or case series involving one subject</td>
</tr>
</tbody>
</table>

Original version:
Development of reporting guidelines

- Comprehensive literature search for:
  - previous guidance documents
  - empirical evidence for poor reporting
  - empirical evidence for bias due to poor reporting

- Workshop of guideline development group
- Comments by additional interested parties
- Invitation at large for comments & criticism (website)
- After publication: revision based on feedback & new evidence on reporting quality
- Transparent decisions (documented in explanatory papers)
Scope of reporting guidelines

- **Not** about methodological quality of studies:

  “Accurate and transparent reporting is like turning the light on before you clean up a room: It (the light) doesn’t clean it for you but does tell you where the problems are.”


- Adherence to reporting guidelines does not guarantee a high-quality study but more transparency on strength & weaknesses
Enforcement or endorsement?

- **RG are guidance - not requirements**
  - Some journals enforce adherence to RG (e.g. by requesting filled checklist from authors)
  - Many journals recommend use of RG in author instructions

- **RG can be useful for different groups:**
  - authors
  - peer reviewers
  - editors
  - readers
Most famous reporting guideline: the CONSORT Statement

= “Consolidated Standards of Reporting Trials“

- First published in 1996; revised in 2001 and 2010
- Set of 25 essential items that should be reported in a trial report
- Flow diagram describing how participants progress during the course of the study
- Comprehensive explanatory paper with examples of good reporting
- www.consort-statement.org
## CONSORT Checklist

<table>
<thead>
<tr>
<th>Section/Topic</th>
<th>Item No</th>
<th>Checklist Item</th>
<th>Reported on page No</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title and abstract</strong></td>
<td>1a</td>
<td>Identification as a randomised trial in the title</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1b</td>
<td>Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts [21,31])</td>
<td></td>
</tr>
<tr>
<td><strong>Introduction</strong></td>
<td>2a</td>
<td>Scientific background and explanation of rationale</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2b</td>
<td>Specific objectives or hypotheses</td>
<td></td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td>3a</td>
<td>Description of trial design (such as parallel, factorial) including allocation ratio</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3b</td>
<td>Important changes to methods after trial commencement (such as eligibility criteria), with reasons</td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>4a</td>
<td>Eligibility criteria for participants</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4b</td>
<td>Settings and locations where the data were collected</td>
<td></td>
</tr>
<tr>
<td>Interventions</td>
<td>5</td>
<td>The interventions for each group, with sufficient details to allow replication, including how and when they were actually administered</td>
<td></td>
</tr>
<tr>
<td><strong>Outcomes</strong></td>
<td>6a</td>
<td>Completely defined prespecified primary and secondary outcome measures, including how and when they were assessed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6b</td>
<td>Any changes to trial outcomes after the trial commenced, with reasons</td>
<td></td>
</tr>
<tr>
<td><strong>Sample size</strong></td>
<td>7a</td>
<td>How sample size was determined</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7b</td>
<td>When applicable, explanation of any interim analyses and stopping guidelines</td>
<td></td>
</tr>
<tr>
<td><strong>Randomisation:</strong></td>
<td>8a</td>
<td>Method used to generate the random allocation sequence</td>
<td></td>
</tr>
<tr>
<td>Sequence generation</td>
<td>8b</td>
<td>Type of randomisation; details of any restriction (such as blocking and block size)</td>
<td></td>
</tr>
<tr>
<td>Allocation concealment mechanism</td>
<td>9</td>
<td>Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned</td>
<td></td>
</tr>
<tr>
<td>Implementation</td>
<td>10</td>
<td>Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions</td>
<td></td>
</tr>
<tr>
<td><strong>Blinding</strong></td>
<td>11a</td>
<td>If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11b</td>
<td>If relevant, description of the similarity of interventions</td>
<td></td>
</tr>
<tr>
<td><strong>Statistical methods</strong></td>
<td>12a</td>
<td>Statistical methods used to compare groups for primary and secondary outcomes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12b</td>
<td>Methods for additional analyses, such as subgroup analyses and adjusted analyses</td>
<td></td>
</tr>
<tr>
<td><strong>Results</strong></td>
<td>13a</td>
<td>For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome</td>
<td></td>
</tr>
<tr>
<td>Participant flow (a diagram is strongly recommended)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Impact of CONSORT (1)

- All leading general medical journals and hundreds of specialist journals support CONSORT

- **Extensions**
  - Cluster trials, non-inferiority and equivalence trials, harms, non-pharmacological treatments, pragmatic trials, abstracts

- **Modifications**
  - REFLECT: RCTs for livestock and food safety

- **Translations in other languages**

- **Other RG are still not widely supported by journals or adhered to by researchers / authors**
Impact of CONSORT (2)

- Some evidence that adoption of CONSORT by journals is associated with improved reporting
  - Review by Plint et al. (Med J Aust 2006)

- Recent empirical study by Hopewell et al. (BMJ 2010)
  - Comparison of RCTs published in 2000 (n=519) and 2006 (n=616)
  - Publication of revised CONSORT in 2001
  - Quality of reporting: proportion of general and methodological CONSORT items reported in publications
  - Improvements for some key items incl. primary outcome, sample size calculation, random sequence generation
  - Remains below acceptable level overall
“Family” of reporting guidelines

STROBE: observational studies
www.strobe-statement.org

STARD: diagnostic studies
www.stard-statement.org

PRISMA: meta-analyses of RCTs

SQUIRE: quality improvement studies
etc.
STROBE Statement (2007)

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies

Erik von Elm, MD; Douglas G. Altman, DSc; Matthias Egger, MD; Stuart J. Pocock, PhD; Peter C. Gøtzsche, MD; and Jan P. Vandenbroucke, MD, for the STROBE Initiative

Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration

Jan P. Vandenbroucke¹, Erik von Elm²,³, Douglas G. Altman⁴, Peter C. Gøtzsche⁵, Cynthia D. Mulrow⁶, Stuart J. Pocock⁷, Charles Poole⁸, James J. Schlesselman⁹, Matthias Egger²,¹⁰ for the STROBE Initiative
EQUATOR Network

= Enhancing the QUALity and Transparency Of health Research

- **Network of reporting initiatives since 2008**
  - ‘Library for Health Research Reporting’ as freely available, up-to-date online resource
  - Education programme (courses for editors, peer reviewers)

- [www.equator-network.org](http://www.equator-network.org)
Welcome to the EQUATOR Network website – the resource centre for good reporting of health research studies

Too often, good research evidence is undermined by poor quality reporting.

The EQUATOR Network is an international initiative that seeks to improve reliability of medical research literature by promoting transparent and accurate reporting of research studies.

Find out how, or get involved.

Highlights

EQUATOR Network at the Peer Review Congress 2009
Wednesday 9th September, Vancouver, Canada.

Prior to the main congress, the EQUATOR Network will run a workshop on major scientific and ethical issues in health research reporting. Dr Richard Horton, Editor-in-Chief of The Lancet, will
How to shift the ‘reporting culture’?

- **Collaboration of parties involved in publication process**
  - Journals: editors, peer reviewers, publishers
  - Researchers / authors
- **But also:**
  - Universities, research institutions, international organisations
  - Funders incl. public and industry
  - Regulatory bodies e.g. ethics committees

- **Working towards:**
  Accurate, complete and transparent reporting of research studies should be considered a ‘norm’

- **How to achieve this?**
  - Provision of tools (e.g. reporting guidelines) and other resources
  - Education and training of groups involved
  - Motivation and incentives for transparent reporting
Good reporting is not an optional extra:

it is an essential component of good research

Thank you!