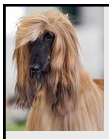


Does peer review work?

(… or is it a shaggy dog story)?

Elizabeth Wager PhD

Publications Consultant
Sideview, Princes Risborough, UK
Visiting Professor, University of Split
Co-Editor-in-Chief *Research Integrity & Peer Review*
liz@sideview.demon.co.uk

Does peer review 'work'?

- Need to define purpose:
  - selection tool,
  - or quality control mechanism?

Jefferson, Wager, Davidoff. *JAMA* 2002;**287**:2786-90

Is peer review a beauty pageant?

or a grooming service?

How do you measure the 'quality' of peer review?

- % acceptance (filter)
- speed
- cost
- consistency (agreement)
- sensitivity to error
- sensitivity to fraud
- selectivity to important articles
- constructiveness (leads to improvements)
- ease of review
- objectivity (bias)

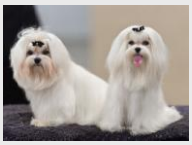
Jefferson, Wager, Davidoff. *JAMA* 2002;**287**:2786-90

Rejection rates

Readership	Journal	Rejection rate
General	*Lancet, NEJM, PLOS Medicine*	>90%
Specialist	*Circulation, Heart*	85% 75%
Sub-specialty	*Jnl of Interventional Cardiology*	50-60%
Super-specialist	*Jnl of Vascular Access*	<60%
Bias to publish	*PLOS One, Drugs in Context*	15-30%

Beauty is subjective …



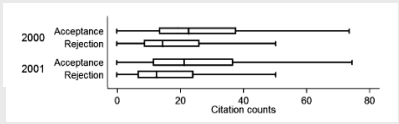




Peer reviewers don't agree much

- Study of 1899 articles
- kappa coefficients 0.10 – 0.21
- 'indicate a low level of agreement between the referees' recommendations concerning acceptance or rejection'

Bornmann & Daniel *Angew Chem* 2008;47:7173-8

Reviewers don't reliably select highly cited articles



Bornmann & Daniel *Angew Chem* 2008;47:7173-8

Journals have different criteria

THE LANCET

"We seek to publish high-quality clinical trials that will alter medical practice"

- *"Our editorial view is that readers can decide for themselves whether or not an article has value or relevance to them, and this is the way that the internet has transformed publication of all kinds. Print publication, because of space limitations, forces decisions on editors based on their judgement of what's of interest to readers. Online publication allows readers to decide what's of interest to them."*

JRSM open

- PLOS | MEDICINE publishes "outstanding research and commentary … We specifically seek to publish articles relevant to clinicians and policymakers across a range of settings that adhere to the highest standards of methodology, ethics and reporting"
- 3% acceptance rate

- PLOS ONE "will rigorously peer-review your submissions and publish all papers that are judged to be technically sound. Judgments about the importance of any particular paper are then made after publication by the readership (who are the most qualified to determine what is of interest to them)"
- 70% acceptance rate

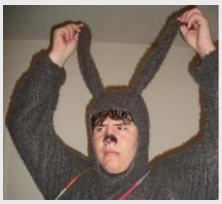
Does peer review detect errors?

Peer review is not effective at detecting errors

- Godlee et al. The impact of blinding and masking on the quality of peer review. *JAMA* 1998;**280**:237-40
- Sent paper (+8 errors) to *BMJ* reviewers

	N*	Mean no. errors identified
Traditional	72	1.9
Open	30	1.8
Masked	59	2.1
Masked + sign	60	1.7

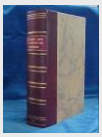
Does peer review prevent fraud?

or falsification?


Rigorous peer-review has not prevented the publication of fraudulent science in major journals

- Hwang Woo-Suk cloning: *Science*
- Jan Hendrik Schön nanotechnology: *Science*
- John Sudbø cancer: *Lancet* & *NEJM*



1785: Literary & Philosophical Society of Manchester

- Review can guarantee only *'the novelty, ingenuity, or importance'* of submissions
- *'Responsibility concerning the truth of facts, the soundness of reasoning … [and] the accuracy of calculations is wholly disclaimed: and must rest alone, on the knowledge, judgement, or ability of the authors who have respectfully furnished such communications'*

Peer review fails to identify good work

- *Nature* publicised the fact that it had published the work of Paul Lauterbur who had just won the Nobel prize (2003)
- Lauterbur politely wrote in to point out that his paper had been published only after he had appealed against a rejection

'Not only does the mechanism of peer review fail to protect us from disasters, in a certain way it guarantees mediocrity: the genius has no peers'

Edsger W Dijkstra

Does peer review improve papers?

trim them?

clean them?

make them conform to standards?

Cochrane reviews

- **Peer review** (submission to acceptance)
- Jefferson *et al*
- *JAMA* 2002;**287**:2784-6
- *update on Cochrane Library (2007)*

- **Technical editing** (acceptance to publication)
- Wager & Middleton
- *JAMA* 2002;**287**:2821-4
- *update on Cochrane Library (2008)*

Cochrane review: limitations

- Focus on medical publications
- Focus on pre-publication, invited peer review
- Strict definition of study quality
- Excludes qualitative research

- Now rather out-of-date

Effectiveness of peer review

- 28 studies
- No clear evidence of effect of reviewer or author concealment (9 studies)
- 2 studies suggest use of checklists may improve quality of reports
- 2 studies showing PR makes papers more readable but ?generalizable
- No effect of: method of contacting reviewers (3), reviewer training (1)

Update on effects of blinding

- 17 studies on effects of blinding
- Failure of blinding (reviewer correctly guessing author identity) 46-73%
- Minimal/no effect on scientific quality of reviewer reports

Hilda Bastian: http://blogs.plos.org/absolutely-maybe/weighing-up-anonymity-and-openness-in-publication-peer-review/

Can blinding reduce bias?

- One RCT found reviewers gave higher ratings to MS by senior cf junior scientists
- AND difference was INCREASED if author identity was removed!
- Another RCT found no difference in acceptance rates for US vs non-US studies
- No clear effect on gender bias (mixed results from studies)

Asking authors to suggest reviewers (WAME survey)

- Some journals do it all the time (eg *BMC*)
- Some journals would *NEVER* ask!

Responded	24 (16 med)
Permit / encourage ANR	14
Use ANR >50% of MS	6
Use ANR 25-50% of MS	8
Add to database	2

Author-nominated reviewers: the evidence

- One study[1] showed author-selected reviewers were slightly less critical
  (mean score 2.51 vs 2.75 where 1=accept, 4=reject)
- Two studies[2,3] show author-selected reviewers perform as well as editor-selected

1 Earnshaw et al. *Ann R Coll Surg* 2000;**82**:133-5

2 Wager et al. *(BMC Medicine* 2006;**4**:13)
3 Schroter et al. *(JAMA* 2006;**295**:314-7)

The Review Quality Instrument

- van Rooyen et al. *J Clin Epidemiol* 1999;**52**:625-9
  BMJ 1999;**318**:23-7
- Rates review according to comments on:
  - importance of research question
  - originality
  - methods strengths & weaknesses
  - presentation
  - interpretation of results
  - specific / constructive suggestions
  - (tone)

RQI contd.

- 7 or 8 questions
- Each scored 1 (=worst) to 5 (=best)
- Overall = mean score (1-5)
- 'Meaningful difference' = 10% (0.4 point)

Are reviewers suggested by authors as good as those chosen by editors?

(1) BMC Study

- Wager, Parkin & Tamber, *BMC Medicine* 2006;**4**:13
- Compared reviews from 100 papers
- No difference in review quality (mean RQI ANR 2.24±0.55 vs ECR 2.34±0.54)
- No difference in tone (2.72 vs 2.82)
- ANRs more likely to recommend acceptance (42 vs 35, p<0.001)

Are reviewers suggested by authors as good as those chosen by editors?

(2) BMJ Study

- Schroter et al. *JAMA* 2006;**295**:314-7
- Compared reviews from 329 papers
- No difference in review quality (mean RQI ANR 2.58 vs ECR 2.64)
- ANRs more likely to recommend acceptance (57% vs 46%)
- ANRs less likely to recommend rejection (13% vs 24%)

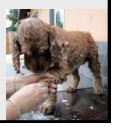
Effectiveness of technical editing

- 32 studies (66 surveys of reference accuracy)
- 'Editorial package' may improve readability and reporting quality
- More intensive editing reduced errors in abstracts and references
- Structuring improved quality of abstracts
- Mixed effects of journal policies / instructions to authors

and peer review is …

- Slow
- Expensive
- Unreliable (would need 6 reviewers for each paper to get statistically reliable result!)
- Subjective (?biased)
- Open to abuse

Wager & Jefferson, *Learned Pub* 2001;**14**:257-63



Journal 'house styles' are rarely 'evidence-based'

Not everybody wants the 'standard' look …

But many manuscripts do need attention

Does peer review 'work'?

- Need to define purpose:
  - selection tool, quality control mechanism
- Some evidence that peer review improves quality of papers
- Not good at detecting major fraud (eg fabrication)

Jefferson, Wager, Davidoff. *JAMA* 2002;**287**:2786-90

What about alternatives to pre-publication, invited review?

Public peer review

Nature trial

- Authors unwilling to have MS posted before publication

MJA trial

- Articles received few comments
- Some articles had no comments
- Comments were superficial

?any studies of

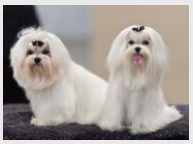
- *F1000 Research* model
- *Frontiers*
- Peerage of Science
- Rubriq (commercial review)
- Preprints – Arxiv, *PeerJ*

Conclusions

- Peer review is not particularly effective at:
  - identifying errors
  - identifying fraud (fabrication / falsification)
  - Difficult to know if it identifies the 'right' articles for a particular journal
- Peer review has many flaws (speed, cost, bias)
- But we don't (yet) have a better system

Thank you …




Backup slides

Does masking raise the quality of reviews?

- Several studies, inconclusive evidence
- Robert & Suzanne Fletcher*: *'journal editors might reasonably choose to blind or not. There appears to be little at stake in their choice'*

*Fletcher R & Fletcher S. Effectiveness of peer review. In *Peer Review in Health Sciences* (2e) Godlee F, Jefferson T (eds), BMJ Books, London, 2003, p.68-9

Effects of open review (i)

- Godlee et al. The impact of blinding and masking on the quality of peer review. *JAMA* 1998;**280**:237-40
- Sent paper (+8 errors) to *BMJ* reviewers

	N*	Mean no. errors identified
Traditional	72	1.9
Open	30	1.8
Masked	59	2.1
Masked + sign	60	1.7

*No difference in response rate

Effects of open review (ii)

- van Rooyen et al. Effect of open peer review on quality of reviews and reviewers' recommendations. *BMJ* 1999;**318**:23-7
- Paired assessments of 125 manuscripts
- Open review increased refusal rate (35% vs 23%)
- No difference in quality between anonymous and signed reviews (3.06 vs 3.09)
- No difference in recommendation
- No difference in speed

Effects of open review (iii)

- van Rooyen et al. *JAMA* 1998;**280**:234-7
- Results from 467 ms (complex design)
- Masking had no effect on quality (both=2.9)
- 33% of reviewers correctly identified authors
- 7% incorrectly identified authors

Conclusions

- Evidence that masked review raises quality or reduces bias is equivocal
- Editors should base decision on knowledge of their own field / area
- It is hard to mask author identity effectively
- Open (signed) review is feasible (for *BMJ*)
- Open review does not affect quality