ed!tage
by CACTUS

Improving peer review performance

Evidence from medical journals

Liz Wager PhD
Sideview

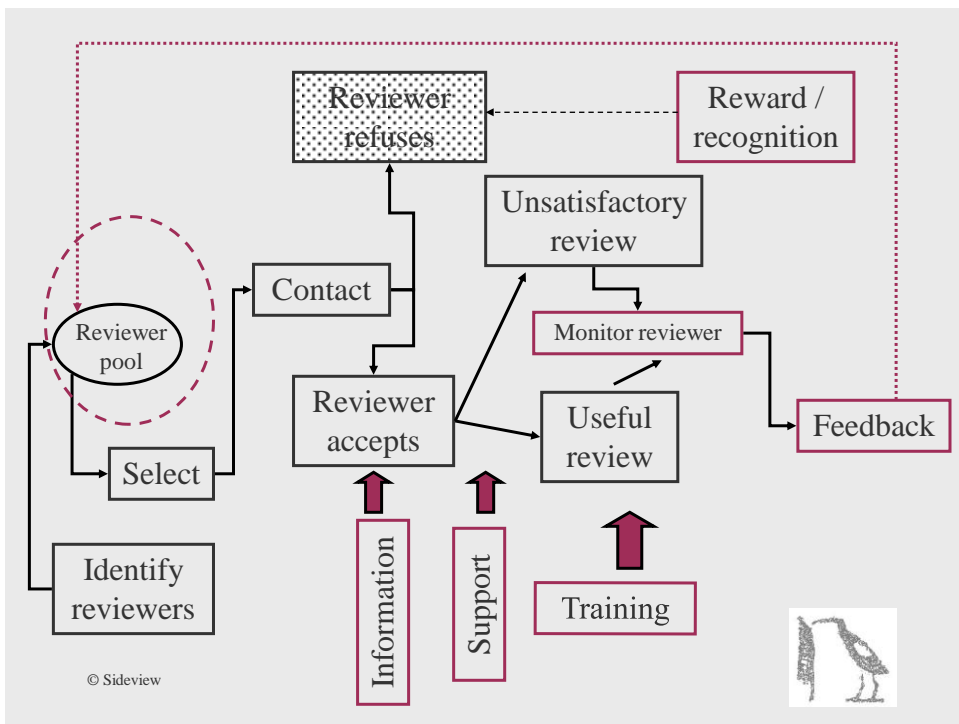# Opportunities for improvement

- Selecting / contacting reviewers
- Guidelines for reviewers
- Systems to reduce bias / prejudice
- Supporting reviewers
- Feedback / sharing reviews
- Monitoring / assessing reviewers
- Training reviewers
- Reviewer rewards / incentives

© Sideview

Identifying reviewers

- Inherited database
- Own contacts
- Author suggestions
- Reviewer / Editorial board suggestions
- Authors of cited papers
- PubMed search

© Sideview

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

© Sideview

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)

© Sideview

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%)

Conclusions

- Author-nominated reviewers (ANRs) produce objective reviews of the same scientific quality as editor-chosen reviewers
- Most journals do not ask for reviewers' views on acceptance (editor decides)
- ANRs less likely to recommend rejection

© Sideview

'Cold calling'

- Many journals identify reviewers from cited references or a PubMed search
- You can often find e-mail addresses from PubMed or Google Scholar
- Will reviewers review for my journal?

© Sideview

Lessons from Croatia

- Editors contacted 236 Croatian doctors
- Would they review for *Lancet* or *CMJ*?
  - Only *CMJ*   72%
  - Only *Lancet* <1% (1 doctor)
  - Yes to both   18%
  - No to both   9%

Marusic et al *CMAJ* 2005;**172**:727

Evidence from India

- *National Medical Journal of India*
- Sent MS to Indian and non-Indian reviewer
- Data on 78 MS (=156 revs), assessed blindly
- Non-Indian reviewers produced higher quality reviews
- Mean score 56.7 vs 48.6 (max 100)

Das Sinha et al. *Natl Med J India* 2000; **13**:105-6

Who are the best reviewers?

- Young (<40)
- Working at good institutions
- Trained in epidemiology and statistics

van Rooyen et al. *JAMA* 1998;**280**:231-3

Stossel. *NEJM* 1985;**312**:658-9

Evans et al. *J Gen Int Med* 1993;**8**:422-8

© Sideview

Does contact method affect response?

- RCT of 'ask first' vs 'just send' (N=283 MS)
- Reply time = shorter for 'ask first'
  (21 vs 25 days, p=0.008)
- *BUT* overall review time = no difference
  (24.7 vs 25.9 days, p=0.19)

Pitkin & Burmeister. *JAMA* 2002;**287**:2795-6

© Sideview

How to contact tardy reviewers?

- RCT phone vs fax vs e-mail (N=378)
- Requested review within 21 days
- Contacted reviewers after 28 days
- Similar numbers returned review within 7 days:
  - phone   68%
  - fax   67%
  - e-mail   67%

Pitkin & Burmeister. *JAMA* 2002;**287**:2794-5

Guidelines for reviewers

- Reviewers are not telepathic
- Journals have different requirements
- It makes sense to give clear guidance
- Guidelines and checklists are appreciated by less experienced reviewers

What might guidelines cover?

- Deadline
- Review method (anonymous, masked, open)
- Competing interests / confidentiality
- Points to consider
  originality, soundness, interpretation, presentation
- Recommendation? (accept, revise, reject)
- Format (length, medium)

Are guidelines effective?

- We know most authors do **not** read them!
  - Pitkin. *NEJM* 1998;339:1006
- Not all reviewers will follow them
- Less experienced reviewers probably appreciate them most
- Guidelines can form the basis for training

*'If you don't have a dream ..
how you gonna have a dream come true?*

Guidance may increase agreement

- J Am Acad Child Adol Psych introduced new (1-10) rating scales + manual
- Asked more concrete questions rather than global judgement
- Reviewer agreement increased from 0.27 to 0.43 (fair to good agreement)

Strayhorn et al. *Am J Psych* 1993;**150**:947-52

Reducing bias and prejudice and protecting reviewers

- Editor should aim to increase objectivity
- Encourage reviewer to criticise the submission **not** the author(s)
- Reviewers may be influenced by: geographic, ethnic, gender, status bias
- Reviewers may also be concerned about future relations with the author(s)

Review systems

	Reviewer knows authors' identity	Authors' identity masked
Reviewer's identity masked	Traditional	Masked (blinded)
Author knows reviewer's identity	Open	?

Reviewer knows authors' identity; review is anonymous

Traditional review

Advantages

- Cheap & simple
- Reviewers can state their views freely

Disadvantages

- Not transparent
- Author may guess reviewer's identity
- Reviewers may be discourteous
- Reviews may be biased

Masked review

Author identity is masked; review is anonymous

Advantages

- May reduce bias
- May increase objectivity
- Reviewers can state their views freely

Disadvantages

- Time-consuming
- Not always possible to mask effectively
- Author may guess reviewer's identity
- Reviews may be discourteous

© Sideview

Open review

Reviewer knows authors' identity; review is signed

Advantages

- Transparent
- No extra cost
- Avoids guessing
- May produce more courteous reviews?

Disadvantages

- Some reviewers may refuse to sign
- Reviewers may be too guarded / not candid

© Sideview

What do most journals do?

- ALPSP survey (200 journals, 40% biomed)
- 60% traditional
- 40% masked
- ? open *(BMJ, JAMA, BMC)*
- 88% anonymous review

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

© Sideview

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

© Sideview

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

Supporting reviewers

- Survey of 733 *JNCI* reviewers
  (66% response, N=481)
- 51% of reviewers do literature search
- Journal supplied abstracts (1 hr/MS)
- 80% found abstracts helpful
- 35% thought abstract affected comments

Hatch & Goodman. *JAMA* 1998;**280**:273-4

Exchanging reviews

- *Natl Med J India* study
- MS sent to Indian and non-Indian reviewer
- Randomized to exchange reviews (or not)
- 38 reviews exchanged, 40 not
- Being told that review would be sent to other reviewer had no effect on review quality
- *But did it affect quality of future reviews?*

Das Sinha et al. *Natl Med J India* 2000; **13**:105-6

Evidence from the US

- *Annals of Emerg Med* randomized reviewers to receive written feedback (copy of other reviewer's review, editor's rating of own review, journal's quality criteria)

- No effect on quality of subsequent reviews for low quality or average quality reviewers

Callaham et al. *JAMA* 2002;**287**:2781-3

Monitoring reviewers

- *'Any jnl that uses 100 or more reviewers probably needs .. a regular system of objective, quantitative rating of review quality by editors'* Mike Callaham

- WAME survey: 42% of jnls rated reviewers
- Overall rating of 1-5 is probably fine
- RQI is well validated (but takes longer)

Training reviewers

- Reviewers appreciate training
- Workshops and distance learning (CD / web) have been offered
- **<u>But</u>** training has never been shown to improve reviewer performance significantly

BMJ study

- 609 *BMJ* reviewers took part (48% response)
- Randomized to attend workshop, receive self-study pack (CD) or control group
- Sent 3 papers containing errors (over 6 months)
- No sustained difference in quality (RQI) or number of errors identified after training

Schroter et al. *BMJ* 2004;**328**:673-5

BMJ study: effect on RQI

	Workshop	Self-study	Control
Review 1	2.72	2.73	2.67
Review 2	2.72	2.85	2.56
Review 3	2.76	2.89	2.74

RQI range 1-5

© Sideview

BMJ study: errors spotted

	Workshop	Self-study	Control
Review 1	2.68	2.68	2.38
Review 2	2.96*	3.14*	2.13
Review 3	3.18	3.37	2.71

Each paper had 9 major errors
* sig diff from control

© Sideview

BMJ study: % advising reject

	Workshop	Self-study	Control
Review 1	70	67	68
Review 2	84	92*	76
Review 3	83	91*	74

*self-study group sig diff from control

BMJ study: conclusions

- Training led to 'slight' improvements
- Did not reach definition of 'editorial significance'
- Self-taught package (CD) seemed more effective than workshop
- CD increased % recommending rejection (which editors considered the correct response)
- Effect of workshop wore off after 6m

Ann Emerg Med study 1

- Non-randomised 'before & after' study
- No measurable effect of 39 reviewers attending 4h workshop on peer review

Callaham et al. *Ann Emerg Med* 1998;**32**:318-22

Ann Emerg Med study 2

- Reviewers randomized to receive detailed feedback (example of good review, editor's rating of own review, copy of other review)
- Mean change 0.06 vs control 0.12 (1-5 scale)
- *'Simple written feedback seems an ineffective educational tool'*

Callaham et al. *JAMA* 2002;**287**:2781-3

Commentary (Frank Davidoff)

- No surprise that 'short, cognitively focused, and largely didactic (passive)' training affects complex skills needed for review
- Discourage 'further use of precious time, energy, and funds for .. intervention that's unlikely to be effective'
- 'broader scientific and scholarly communities' should take responsibility for producing and rewarding good reviewers

Davidoff *BMJ* 2004;**328**:657-8

Rewards & incentives

- Does payment:
  - reduce refusals?
  - accelerate reviews?
  - raise / lower review standard?
- How else can we recognise reviewers?
- Can we increase academic recognition?

Summary & conclusions (1)

- Consider several sources for enlarging reviewer database, e.g. author-nominated reviewers
- Try to use reviewers from outside your own country
- Young reviewers are probably best!

© Sideview

Summary & conclusions (2)

- Phoning reviewers first doesn't gain much
- Phone, fax, e-mail equally effective methods for chasing tardy reviewers
- Guidelines may improve consistency
- Review systems should ensure objectivity
- Evidence re masking reviews is unclear

© Sideview

ed!tage
by CACTUS

Summary & conclusions (3)

- Supplying abstracts may help reviewers
  (but ? who will do this)
- Consider monitoring reviewer performance
- Exchanging reviews does not affect quality
- No evidence that workshops improve
  performance
- Self-study (distance learning) may be
  effective

© Sideview

---

ed!tage
by CACTUS

Suggestions

- Skills needed for critical evaluation
  (evidence-based medicine) are the same as
  those needed to review
- Work with local universities (medical
  schools) to improve training in critical
  evaluation
- Liaise with universities to increase
  recognition of role of reviewing

© Sideview

Suggestions (2)

- Could use similar model to train potential authors
- Work with medical schools / continuing education centres to devise courses on writing, stats, peer-review process
- Already have system to reward publication!

© Sideview

Hope is the companion of power and the mother of success. For those of us who hope strongest have within us the gift of miracles.
Sydney Bremer

Everyone is trying to accomplish something big, not realizing that life is made up of little things.
Frank A. Clark

© Sideview