

## Bias Methods Group meeting

This short presentation is about biases in non-randomised studies (NRS)

I accept as given, that

- all authors (of primary studies and reviews) have an “agenda” (bias)
- RCTs offer the best framework to protect against authors’ biases
- This framework is to do with much more than just randomisation

I am sceptical about

- the value to decision-makers of most NRS of effectiveness (very much more susceptible to ‘standard’, publication and reporting biases)
- the ability of most review authors to appraise the risk of bias in primary NRS

I worry that

- most users of evidence have a prior point of view, i.e. are seeking for evidence to justify an argument rather than to inform a choice
- the remainder do not have a prior point of view but are seeking to make a decision based on ‘evidence’ so that they can justify their actions.

Aim: to illustrate the complexity of appraising a reasonably good NRS  
(prospective cohort study, based on a protocol)

- P** people presenting with an **aneurismal subarachnoid haemorrhage** (SAH)
- I** repair of the aneurysm by **endovascular 'coiling'**
- C** repair of the aneurysm by **surgical 'clipping'**
- O** **death or dependency** (composite)

## International Sub-Arachnoid Haemorrhage Trial (ISAT)

- multicentre, international RCT
- recruited from 1994 (pilot) to May 2002); stopped early following interim analysis
- final result: coiling vs. clipping, 23.5% vs 30.9% dead or dependent at 1 year; ARR 7.4% (3.6 to 11.2); risk ratio = 0.76 (0.66 to 0.87)
- equivalent to clipping vs. coiling, odds ratio = **1.45 (1.19 to 1.77)**

## [UK] National Study of Sub-Arachnoid Haemorrhage (NSSAH)

- multicentre, UK cohort study
- only about half of all centres had staff and equipment to do repairs by coiling
- no a priori objective to compare coiling and clipping (set up as 'comparative audit')
- recruited from Sep 2001 to Sep 2002 (objective: all patients in 1 year)
- final 'adjusted' result: clipping vs. coiling, odds ratio = **1.17 (0.87 to 1.57)**

The NSSAH result was publicised by neurosurgeons in the British Journal of Neurosurgery – claiming that the NSSAH reported a **more appropriate** **“effectiveness” comparison** between the repair techniques (cf. efficacy).

NSSAH Steering Committee responded pointing out:

- (a) that the two studies were not directly comparable, and
- (b) there are multiple explanations for the observed ‘discrepancy’

[Reeves et al. Br J Neurosurg 2007;21:318-327]

(The ISAT investigators also replied separately.)

Reasons why the two studies were not directly comparable:

1. The ISAT was analysed according to the principle of **intention-to-treat**. NSSAH could **not** be analysed in this way because only the actual mode of repair, not the intention, was recorded. In the ISAT, 4 times more patients crossed over from clipping to coiling than from coiling to clipping – a similar phenomenon is likely to have occurred in NSSAH but was not recorded.
2. ISAT included 4.9% patients who were randomised but who did not have a repair, e.g because of neurological deterioration before the repair could be done. In the NSSAH, 9.3% of patients had no repair and were excluded because they could **not** be allocated [by intention-to-treat] to group.
3. ISAT had missing data at 1 year for only 1.1% of randomised patients. The NSSAH had missing data for 9.4% of patients who had a repair.
4. Times-to-repair for all ISAT participants could be derived with reference to the time of randomisation (repairs were done more quickly with coiling). In the NSSAH, times from bleed-to-repair were known and did not differ by mode of repair; however, these times took no account of transfers between centres, almost all of which were from a centre that did not have a coiling facility to one that did.

Multiple explanations for the observed 'discrepancy':

5. Chance (there may be no real discrepancy; 95% CIs for estimates substantially overlap) [LIKELY]
6. Residual confounding of the effect estimate observed in the NSSAH [LIKELY]
7. Differences in study design (as described at 1-4) [LIKELY]
8. Differences in the patients included in the NSSAH who were not randomised in the ISAT [LIKELY, NOT SUCH THAT OBSERVATIONAL COMPARISON SHOULD FAVOUR SURGERY]
9. Differences in the way the interventions were provided in UK centres (the majority) that did not participate in the ISAT [LIKELY]

This example arose (and the BJNS paper was written ) because a surgical member of the NSSAH research team wanted the team to support the third party comment on the NSSAH and the chair of the steering committee disagreed.

What are the implications for the Cochrane Collaboration? Consider a review team doing a review of coiling versus clipping:

- With a more ‘relaxed’ Collaboration policy about including non-randomized studies, would the review team want to include the NSSAH to “complement” the single RCT?
- Would the lack of comparability of the studies been apparent to reviewers?
- How would the ‘discrepancy’ [or similarity] between effect estimates have been interpreted without the commentary by the NSSAH researchers?
- Given the claims of surgeons about limitations of the ISAT, might the NSSAH be considered ‘better’ evidence for reasons of applicability?
- Might / should inclusion of the NSSAH lead to a down-weighting of the evidence from the ISAT?