Title **Evaluating Nonrandomized Intervention Studies**

Agency **NCCHTA, National Coordinating Centre for Health Technology Assessment**
Mailpoint 728, Boldrewood, University of Southampton, Southampton SO16 7PX, United Kingdom;
Tel: +44 2380 595586, Fax: +44 2380 595639

Reference Health Technol Assess 2003;7(27). Sept 2003. www.ncchta.org/execsumm/summ727.htm

Aim

To consider methods and related evidence for evaluating bias in nonrandomized intervention studies.

Conclusions and results

In the systematic reviews, 8 studies compared results of randomized and nonrandomized studies across multiple interventions using meta-epidemiological techniques. We identified 194 tools for assessing nonrandomized studies: 60 tools covered at least 5 of 6 prespecified internal validity domains; 14 tools covered 3 of 4 core items of particular importance for nonrandomized studies; 6 tools were thought suitable for systematic reviews. Of 511 systematic reviews that included nonrandomized studies, only 169 (33%) assessed study quality, and 69 reviews investigated the impact of quality on study results in a quantitative manner. The new empirical studies estimated the bias associated with nonrandom allocation, and found that the bias could lead to consistently over- or underestimating treatment effects. Also, bias increased variation in results for both historical and concurrent controls. The biases were large enough to lead studies falsely to conclude significant findings of benefit or harm. Four strategies for case-mix adjustment were evaluated: none adequately adjusted for bias in historically and concurrently controlled studies. Logistic regression on average increased bias. Propensity score methods performed better, but were not satisfactory in most situations. Investigation revealed that adjustment can only be adequate in the unrealistic situation when selection depends on a single factor.

Recommendations

Results of nonrandomized studies may differ from results of randomised studies of the same intervention and may be misleading when the treated and control groups are similar in key prognostic factors. Standard methods of case-mix adjustment do not guarantee removal of bias. Residual confounding may be high even with prognostic data, and adjusted results might appear more biased than unadjusted results. Most quality assessment tools omit key quality domains. Healthcare policies based on nonrandomized studies or systematic reviews of nonrandomized studies may need re-evaluation if the uncertainty in the true evidence base was not fully appreciated when policies were made. Nonrandomized studies should only be undertaken when RCTs are infeasible or unethical.

Methods

Three systematic reviews and new empirical investigations were conducted. The reviews considered, regarding nonrandomized studies:

1. Existing evidence of bias
2. Content of quality assessment tools, and
3. Ways that study quality has been assessed and addressed.

The empirical investigations were conducted, generating nonrandomized studies from two large, multicenter randomized controlled trials (RCTs) and selectively resampling trial participants according to allocated treatment, centre, and period.

Further research/reviews required

Apply resampling methodology in other clinical areas to ascertain whether the biases described are typical. Develop or refine existing quality assessment tools for nonrandomized studies. Investigate how quality assessments of nonrandomized studies can be incorporated into reviews and the implications of individual quality features for interpretation of a review's results. Examine reasons for the apparent failure of case-mix adjustment methods. Further evaluate the role of the propensity score.

Written by Mr Jonathan Deeks, Centre for Statistics in Medicine, University of Oxford, UK