Centre for Reviews and Dissemination

THE UNIVERSITY of York

Comparison of adverse effects data derived from different study designs

Su Golder[1], Yoon K Loke[2]

[1]Centre for Reviews and Dissemination, University of York, [2]School of Medicine, University of East Anglia (su.golder@york.ac.uk)

Introduction

There is considerable debate about the relative utility of different study designs in generating reliable quantitative estimates of risk of adverse effects. A range of study designs, including randomised controlled trials (RCTs) and non-randomised studies such as cohort or case-control studies may potentially record adverse effects of interventions, and provide useful data for systematic reviews and meta-analysis. However, there are strengths and weaknesses inherent to each study design, and different estimates and inferences about adverse effects may arise depending on study type.

This research aimed to systematically review meta-analyses or methodological studies which have compared estimates of harm (for specific adverse effects) reported in one study design with those reported in another study design for the same intervention and adverse effect.

Methods

Studies were identified by searching 10 databases, handsearching key journals, conference proceedings and websites and by reference checking, citation searching and contacting experts. Studies were included where a pooled relative measure of an adverse effect (odds ratio or risk ratio) from one study design could be directly compared, using the ratio of odds ratios (RORs), with the pooled estimate for the same adverse effect from another study design.

Results

Forty studies, yielding 160 pooled sets of comparisons were identified for inclusion. In almost all instances, the estimates of harm obtained from meta-analyses of the different study designs being compared had 95% confidence intervals that overlapped. Nearly two-thirds of these gave results from which similar conclusions would be drawn (both types of study showing a significant increase or significant decrease or no significant difference in harms associated with the intervention under investigation) (Table 1).

Table 1. Confidence interval overlap and agreement between study designs

Study design comparisons	Confidence interval overlap	Agreement in findings between the study designs			Discrepancy in findings between the study designs		
		Both showed a significant increase	Both did not identify any significant difference	Both showed a significant decrease	Significant risk increase in one vs. significant risk decrease in the other	Significant increase in one vs. no significant difference in the other	Significant decrease in one vs. no difference in the other
RCTs vs observational studies (N=32)	29 (91%)	6 (19%)	13 (41%)	3 (9%)	1 (3%)	8 (25%)	1 (3%)
RCTs vs cohort studies (N=16)	16 (100%)	3 (19%)	8 (50%)	0	0	5 (31%)	0
RCTs vs case-control studies (N=10)	9 (90%)	2 (20%)	2 (20%)	0	0	6 (60%)	0
Cohort vs case-control studies (n=64)	60 (94%)	19 (27%)	23 (38%)	0	1 (2%)	20 (31%)	1 (2%)
Cohort vs cross-sectional studies (n=18)	18 (100%)	4 (22%)	11 (61%)	0	0	3 (17%)	0
Cohort vs ecological studies (n=1)	1 (100%)	0	1 (100%)	0	0	0	0
Case-control vs cross sectional studies (n=18)	18 (100%)	4 (22%)	11 (61%)	0	0	3 (17%)	0
Cross-sectional vs ecological studies (n=1)	1 (100%)	0	1 (100%)	0	0	0	0

In only two meta-analysis was there opposing direction of effect that was statistically significant. In these instances whilst one study design identified a protective effect another type of study design identified an increased risk of the outcome, both these meta-analysis involved menopausal hormone therapy.

Discussion

Most pooled results from the different study designs concurred in terms of identifying a significant increase or decrease, or no significant difference in risk of adverse effects. Where there was discrepancy, the difference was usually a finding of no significant risk of adverse effects with one study design, in contrast to a significant risk of adverse effects from the other study design. This may reflect the limited size of the included studies to identify significant differences in rare adverse effects.

The ratio of risk ratios did not suggest any consistent differences from meta-analysis of different study designs. While there are a few instances of sizeable discrepancies, the pooled estimates indicate that in the scheme of things (particularly where larger, more precise primary studies are available), meta-analysis of one study design should yield adverse effects estimates that broadly match meta-analysis of another study design.

The pooled ratio of odds ratios of different types of study designs indicated no statistically significant difference in estimates of adverse effects derived from different study designs (Table 2).

Table 2. Ratio of odds ratios of adverse effects in study design comparisons

Study design comparison	Pooled ratio of odds ratios (RORs)	Heterogeneity
RCTs versus cohort studies (n=16)	1.02 (95% CI 0.82 to 1.28)	I^2=43%
RCTs versus case-control studies (n=10)	0.84 (95% CI 0.57 to 1.23)	I^2=54%
RCTs versus studies described as 'observational' (n=32)	1.08 (95% CI 0.94 to 1.22)	I^2=60%
Cohort studies versus case-control studies (n=64)	0.94 (95% CI 0.87 to 1.01)	I^2=55%
Cohort studies versus cross-sectional studies (n=18)	0.97 (95% CI 0.89 to 1.07)	I^2=10%
Case-control studies versus cross-sectional studies (n=18)	1.07 (95% CI 0.95 to 1.21)	I^2=26%

Conclusions

These findings have important implications for the conduct of systematic reviews of harm. Although there are strengths and weaknesses of each study design, the empirical evidence from this overview indicates that there is, in general, no difference between estimates on the risk of adverse effects obtained from meta-analyses of the different study designs compared. Instead of restricting the analysis to certain study designs, it may be preferable for systematic reviewers of adverse effects to evaluate a broad range of studies that can help build a complete picture of any potential harm and improve the generalisability of the review without loss of validity.

This research was undertaken by Su Golder as part of an MRC fellowship. The views expressed in this presentation are those of the authors and not necessarily those of the MRC.