Agreement between randomized and non-randomized studies: the effects of bias and confounding

Henry, David A; Moxey, Annette J.; O'Connell, Dianne

Published: 01/10/2001

Document Version:
Peer reviewed version

Agreement between randomized and non-randomized studies - the effects of bias and confounding

Contact Author:

David Henry
School of Population Health Sciences
Faculty of Medicine and Health Sciences
Discipline of Clinical Pharmacology
The University of Newcastle, NSW, Australia.

Address:
Level 5, Clinical Sciences Building
Mater Misericordiae Hospital
Waratah, NSW 2298, Australia

Telephone:
+61 2 49211856

Fax:
+61 2 49602088

Email:
mddah@mail.newcastle.edu.au

Agreement between randomized and non-randomized studies - the effects of bias and confounding

David Henry, Annette Moxey, Dianne O'Connell. School of Population Health Sciences, Faculty of Medicine and Health Sciences, The University of Newcastle, Australia.

Background: Formal comparisons of the results of randomized and non-randomized studies have led to conflicting conclusions. The number of potential topics for study is huge, and levels of agreement vary with selection of interventions and clinical settings.
Objective: To determine the extent to which some important sources of bias affecting randomized trials (RCT) and observational studies (OS) influence the levels of agreement between these designs.
Methods: We performed systematic reviews of RCTs and OS for 7 intervention/outcome pairs chosen because of the likelihood of bias and confounding. These were: interventions to minimize the need for allogeneic blood transfusion (unblinded, poor randomization, subjective outcomes); the impact of laparoscopic cholecystectomy (lap chole), compared with open or mini-laparotomy, on post-operative infections and bile duct injury (variable quality trials, operator skill dependent); the impact of antioxidants on death from malignancy and cardiovascular disease (different interventions, healthy cohort effect); and the effects of hormone replacement therapy (HRT) on cardiovascular and overall mortality (healthy cohort effect). Articles identified through electronic and bibliographic searches were reviewed independently by two raters. Adjusted and unadjusted RRs were pooled using inverse variance weights, and Metaview 4.1 was used to pool crude RRs. We assessed qualitative agreement as +/+ when RCTs and OS (respectively) agreed on the direction of a statistically significant effect, -/- when there was agreement on the absence of such an effect, and +/- or -/+ when there was disagreement We classified quantitative agreement as 0.1or 0.2 if pooled estimates of RR differed by no more than these values, and >0.2 if they did.
Results: The greatest level of agreement between RCTs and OS was seen with blood-sparing techniques (4 comparisons, 59 RCTs and 104 OS) with agreement ranging from +/+ 0.1 to +/+ 0.2. Antioxidants (4 comparisons, 9 RCTs and 13 OS) gave mixed results with agreement graded as -/- >0.2 and -/- 0.2 for CVD mortality with beta-carotene and lung cancer with Vit E, and -/+ >0.2 for CVD mortality with Vit E and lung cancer with beta-carotene. For HRT (2 comparisons, 3 RCTs and 15 OS) agreement was poor for all cause mortality and CVD mortality (-/+ >0.2 for both). Likewise, the agreement was poor for lap chole (2 comparisons, 11 RCTs, 29 OS): -/+ >0.2 for both outcomes.
Conclusions: In this series RCTs and OS studies agreed when the RCTs were of poor quality and evaluated subjective outcome measures. The high quality surgical trials had results closer to the null, but were of insufficient size to quantify adverse effects. Healthy cohort effects probably explain the discrepant findings between RCTs and OS of antioxidants and HRT. To some extent discrepancies between the results of randomized and non-randomized studies can be anticipated from a knowledge of likely sources of bias and confounding. However agreement between the study designs may indicate that they are both giving inaccurate results.

[460 words]

Agreement between randomized and non-randomized studies - the effects of bias and confounding

Background: Formal comparisons of the results of randomized and non-randomized studies have led to conflicting conclusions. The number of potential topics for study is huge, and levels of agreement vary with selection of interventions and clinical settings.
Objective: To determine the extent to which some important sources of bias affecting randomized trials (RCT) and observational studies (OS) influence the levels of agreement between these designs.
Methods: We performed systematic reviews of RCTs and OS for 7 intervention/outcome pairs chosen because of the likelihood of bias and confounding. These were: interventions to minimize the need for allogeneic blood transfusion (unblinded, poor randomization, subjective outcomes); the impact of laparoscopic cholecystectomy (lap chole), compared with open or mini-laparotomy, on post-operative infections and bile duct injury (variable quality trials, operator skill dependent); the impact of antioxidants on death from malignancy and cardiovascular disease (different interventions, healthy cohort effect); and the effects of hormone replacement therapy (HRT) on cardiovascular and overall mortality (healthy cohort effect). Articles identified through electronic and bibliographic searches were reviewed independently by two raters. Adjusted and unadjusted RRs were pooled using inverse variance weights, and Metaview 4.1 was used to pool crude RRs. We assessed qualitative agreement as +/+ when RCTs and OS (respectively) agreed on the direction of a statistically significant effect, -/- when there was agreement on the absence of such an effect, and +/- or -/+ when there was disagreement We classified quantitative agreement as 0.1or 0.2 if pooled estimates of RR differed by no more than these values, and >0.2 if they did.
Results: The greatest level of agreement between RCTs and OS was seen with blood-sparing techniques (4 comparisons, 59 RCTs and 104 OS) with agreement ranging from +/+ 0.1 to +/+ 0.2. Antioxidants (4 comparisons, 9 RCTs and 13 OS) gave mixed results with agreement graded as -/- >0.2 and -/- 0.2 for CVD mortality with beta-carotene and lung cancer with Vit E, and -/+ >0.2 for CVD mortality with Vit E and lung cancer with beta-carotene. For HRT (2 comparisons, 3 RCTs and 15 OS) agreement was poor for all cause mortality and CVD mortality (-/+ >0.2 for both). Likewise, the agreement was poor for lap chole (2 comparisons, 11 RCTs, 29 OS): -/+ >0.2 for both outcomes.
Conclusions: In this series RCTs and OS studies agreed when the RCTs were of poor quality and evaluated subjective outcome measures. The high quality surgical trials had results closer to the null, but were of insufficient size to quantify adverse effects. Healthy cohort effects probably explain the discrepant findings between RCTs and OS of antioxidants and HRT. To some extent discrepancies between the results of randomized and non-randomized studies can be anticipated from a knowledge of likely sources of bias and confounding. However agreement between the study designs may indicate that they are both giving inaccurate results.

[460 words]