Reader's guide to critical appraisal of cohort studies: 1. Role and design.

Paula A. Rochon
Baycrest Centre for Geriatric Care

Jerry H. Gurwitz
University of Massachusetts Medical School

Kathy Sykora
University of Toronto

See next page for additional authors

Follow this and additional works at: https://escholarship.umassmed.edu/meyers_pp

Part of the Health Services Research Commons

Repository Citation
https://escholarship.umassmed.edu/meyers_pp/62

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in Meyers Primary Care Institute Publications and Presentations by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.
Reader's guide to critical appraisal of cohort studies: 1. Role and design.

Authors
Paula A. Rochon, Jerry H. Gurwitz, Kathy Sykora, Muhammad M. Mamdani, David L. Streiner, Susan Garfinkel, Sharon-Lise T. Normand, and Geoffrey M. Anderson
Randomised trial or cohort study?

Cohort studies are similar to randomised controlled trials in that they compare outcomes in groups that did and did not receive an intervention. The main difference is that allocation of individuals is not by chance. Table 1 gives some important similarities and differences between the two types of study. Because they are expensive and recruiting patients can be difficult, randomised controlled trials are generally short term and used to determine efficacy in selected populations under strict conditions. Cohort studies can be used to determine if the efficacy observed in randomised trials translates into effectiveness in broader populations and more realistic settings and to provide information on adverse events and risks.5

Selection bias as a threat to validity

The internal validity of a study is defined as the extent to which the observed difference in outcomes between the two comparison groups can be attributed to the intervention rather than other factors. The biggest advantage of randomised controlled trials compared with cohort studies is the random allocation process enhances the internal validity of a study by minimising selection bias and confounding. This paper relies on the definitions provided by CONSORT (box 1).7

Allocation by chance in a randomised controlled trial should mean that the groups being compared are similar in terms of both measured and unmeasured baseline factors.5 This is not so in cohort studies, and therefore cohort studies are vulnerable to selection bias. In cohort studies, factors that determined whether a person received the intervention could result in the groups differing in factors related to the outcome, either because people were preferentially selected to receive one treatment or because of choices that they made. These baseline differences in prognosis could confound the assessment of the effect of the intervention.8

In cohort studies care must be taken to minimise, assess, and deal with selection bias. A comprehensive...
approach is needed that includes the selection of appropriate comparison groups, the identification and assessment of the comparability of potential confounders between those comparison groups, and the use of sophisticated statistical techniques in the analysis.

### Comparison groups in cohort studies

The essence of any cohort study is the comparison of outcomes between people who received the intervention and those who did not. For example, to answer the question, “Do patients who receive an atypical antipsychotic drug have an increased risk of hip fracture?” a cohort study must ask: “What would have happened to these patients if they had not received the atypical antipsychotic drug?”

Ideally, the comparison group in the cohort study should be identical to the intervention group, apart from the fact that they did not receive the intervention. This ideal comparison group is described by methodologists as providing the “counterfactual” or “potential outcome.” In reality, this ideal comparison group does not exist. Part of the art of designing a cohort study is choosing comparison groups that approach this ideal in order to minimise selection bias while maintaining clinically relevance.

The analysis of the association between antipsychotic drugs and hip fracture can be used to define the types of comparisons that could be found in cohort studies. For any specific intervention (such as exposure to atypical antipsychotics) two factors—the exposure and the outcome—can be measured:

1. Selection bias—systematic error in creating intervention groups, causing them to differ with respect to prognosis. The groups differ in measured or unmeasured baseline characteristics because of the way in which participants were selected for the study or assigned to their study groups.
2. Confounding—a situation in which the estimated intervention effect is biased because of some difference between the comparison groups apart from the planned interventions such as baseline characteristics, prognostic factors, or concomitant interventions. For a factor to be a confounder, it must differ between the comparison groups and predict the outcome of interest.

### Questions to ask when assessing a cohort study design

#### What comparison is being made?

Published studies may include more than one type of comparison, but the focus of any appraisal of a cohort study is on an individual comparison between an intervention group and a comparison group in a defined population. A well written study should contain a clear definition of why the two groups were selected and how they were defined. This information is essential for assessment of clinical relevance and potential for selection bias.

#### Does the comparison make clinical sense?

The clinical relevance of comparisons needs to be assessed for each case. In the analysis of antipsychotic use and hip fracture, for instance, all four types of comparison might be relevant. However, this might not be true in other analyses. For example, although it would be possible for a cohort study to compare HIV positive patients receiving antiretroviral therapy with those receiving no intervention, this comparison would be irrelevant to many clinicians. A more relevant cohort study would compare patients receiving one antiretroviral therapy with patients receiving another intervention. In contrast, a clinically relevant study of the adverse effects of a commonly used treatment such as a non-steroidal anti-inflammatory drug might include a comparison with a no intervention population since no drug treatment could be a realistic option for some people.

Cohort studies should not only describe the populations being compared but also include a discussion of the clinical context for that comparison and provide a justification for the comparison. Readers of these studies should determine if the study makes a comparison that is realistic and relevant to their decision needs.

---

**Table 1** Comparison of cohort studies and randomised controlled trials

<table>
<thead>
<tr>
<th>Item</th>
<th>Cohort studies</th>
<th>Randomised controlled trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Populations studied</td>
<td>Diverse populations of patients who are observed in a range of settings</td>
<td>Highly selected populations recruited on the basis of detailed criteria and treated at selected sites</td>
</tr>
<tr>
<td>Allocation to the intervention</td>
<td>Based on decisions made by providers or patients</td>
<td>Based on chance and controlled by investigators</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Can be defined after the intervention and can include rare or unexpected events</td>
<td>Primary outcomes are determined before patients are entered into study and are focused on predicted benefits and risks</td>
</tr>
<tr>
<td>Follow-up</td>
<td>Many cohort studies rely on existing experience (retrospective studies) and can provide an opportunity for long follow-up</td>
<td>Prospective studies; often have short follow-up because of costs and pressure to produce timely evidence</td>
</tr>
</tbody>
</table>

**Box 1: CONSORT definitions of selection bias and confounding**

- **Selection bias**—a systematic error in creating intervention groups, causing them to differ with respect to prognosis. The groups differ in measured or unmeasured baseline characteristics because of the way in which participants were selected for the study or assigned to their study groups.
- **Confounding**—a situation in which the estimated intervention effect is biased because of some difference between the comparison groups apart from the planned interventions such as baseline characteristics, prognostic factors, or concomitant interventions. For a factor to be a confounder, it must differ between the comparison groups and predict the outcome of interest.

**Box 2: Possible types of comparisons in cohort study**

1. **General population**
   - 1 Intervention vs alternative intervention
   - 2 Intervention vs no intervention
2. **Restricted population**
   - 3 Intervention vs alternative intervention
   - 4 Intervention vs no intervention
Selection bias occurs when there is something inherently different between the groups being compared that could explain differences in the observed outcomes. One powerful strategy to minimise selection bias is to restrict inclusion in the study to those with a defined diagnosis or specific characteristics. Restricting the groups to a specific characteristic removes the potential for bias related to that characteristic and can reduce differences in related characteristics. Table 2 presents data from a cohort of older adults given atypical antipsychotics and a no intervention comparison group. Patients taking atypical antipsychotics were over 12 times more likely (63.1% vs 4.7%) to have dementia. Dementia is related to the risk of hip fracture, and this imbalance may be an important source of confounding. Restricting the study to people with dementia eliminates this source of confounding and reduces selection related to age as the mean age difference between the groups dropped from years to months.

An inevitable consequence of restriction is reduced sample size. In the example, the sample decreased from 1.3 million to about 80,000 when the dementia restriction was applied. When smaller databases are being used, restriction can greatly limit the power of the study. Restriction on the basis of clinical characteristics limits the generalisability of the findings. The more restrictive the population, the less generalisable the results.

It is important to keep in mind the effect the choice of comparison groups will have on potential selection bias when evaluating a cohort study. Some sources of selection bias are clear—for example, if access to atypical antipsychotics was limited to patients of specialists this could result in patients who received these drugs being different from those who did not. Some sources of bias may be more subtle. For example, if doctors thought that atypical antipsychotics had fewer side effects than typical antipsychotics, they might preferentially use the atypical antipsychotics in frailer patients. This form of selection bias, referred to as channelling bias or confounding by indication, occurs when patients are assigned to one intervention or another on the basis of prognostic factors and is key issue in cohort studies.

Readers should recognise the potential for selection bias in all cohort studies and carefully consider possible sources of bias. In the next article we will outline the link between selection bias and confounding and describe a strategy for identifying and assessing the potential for confounding.

We thank Andreas Laupacis for his comments and Jennifer Gold, Michelle Luxer, and Monica Lee for help in preparing the manuscript.

Contributors and sources: The series is based on discussions that took place at regular meetings of the Canadian Institute for Health Research chronic disease new emerging team. PAR is a geriatrician with extensive research experience in cohort studies of prescription drugs who wrote the first draft of this article and is the guarantor. JHG and MM are clinicians and researchers and SELIN and DLS are statisticians who commented on drafts of this paper. KS programmed and conducted analyses and SF conducted literature searches and reviews. PAR and GMA conceived the idea for the series and GMA worked on drafts of this article and coordinated the development of the series.

Funding: This work was supported by a CIHR operating grant (CIHR No. MOP 533124) and a CIHR chronic disease new emerging team programme (NET-54010).

Competing interests: None declared.

1 Garvie JH, Col NF, Avorn J. The exclusion of the elderly and women from clinical trials in acute myocardial infarction. JAMA 1992;268:1417-22.

(accepted 18 February 2005)