Author's response to reviews

Title: Measures and models for causal inference in cross-sectional studies: arguments for the appropriateness of the prevalence odds ratio and related logistic regression

Authors:

Michael E Reichenheim (michael@ims.uerj.br)
Evandro S F Coutinho (esfcoutinho@ensp.fiocruz.br)

Version: 3 Date: 4 July 2010

Author's response to reviews: see over
Dear Editor(s),

We would like to thank both reviewers for their assessments and the Journal for the opportunity of re-submitting the paper “Measures and models for causal inference in cross-sectional studies: arguments for the appropriateness of the prevalence odds ratio and related logistic regression” to BMC-Medical Research Methodology. As requested, below are our point-by-point responses to the reviewers’ concerns (in blue). Please, note that all changes in the 2nd version of the manuscript are also coloured blue.

Kind regards,
Michael E. Reichenheim (p/ Evandro S. F. Coutinho)

Reviewer #1 (RG)

1. Major Compulsory Revisions

1.1 The Abstract has a background, discussion and summary. It is usual to have a methods and results section. The first sentence looks like methods, and the remainder results. The discussion and summary sections should be one or the other. The background has a section called discussion, and this heading should be deleted.

This is because we submitted the paper for the ‘Debate’ article type which requires this structure according to the BMC Medical Research Methodology ‘Instructions for authors’. Thus, we opted to keep it this way.

1.2 The paper uses the notation and terms of reference 9, but it would be useful to include the definitions of the many terms used. For example CI: proportion who develop the disease in time t1 – t0 and prevalence: probability of being a case at time t. It is not clear whether point or period prevalence is used.

Although this could certainly help some, we believe that most readers attracted to this kind of paper would know the definitions asked for by the reviewer. In the interest of keeping the paper as short as possible we plead not to include these definitions. However, we now explicitly write that we
mean point prevalence when referring to this measure in the section on ‘Formal relations between measures’ (page 7).

1.3 The equations 1 to 9 are based on many assumptions (as given in reference 9), and these are stated in the previous section. It would seem that they should be in the section that derived the equations.

Following the reviewer’s suggestion we moved the last two paragraphs of the section ‘Purposes of a cross-sectional studies’ to the next (‘Relating prevalence and incidence ratio measures’), but opted to place them into a new subsection named ‘Structuring conditions’. We slightly changed the first paragraph of Discussion to accommodate this change.

1.4 The paper uses the terms ‘simulations’ but they are ‘scenarios’ based on the equations. There are no simulations in the sense that data are generated using a simulation procedure. I suggest that these be changed throughout the paper.

The terms ‘simulated’ and ‘simulation’ were all substituted.

1.5 The scenarios assume that the ratio of prevalence for the 2 groups is constant and equal to 2. This should be justified, as my preference would be to make the ratio of incidence density equal to 2. In survival analyses, the hazard function is modelled in terms of the predictors, and the incidence density is similar.

We request to keep as is since this is actually our main perspective, i.e., that given data arising from cross-sectional studies with a certain ‘constant’ value, one wants to know what they represent in terms of the incidence measures. Reversing this would completely break our line of thought. We used 2.0 since this is a simple and quite common value found in surveys.

As for the justification, we write the following at the start of section ‘Exploring several scenarios’:

“The scenarios that follow assume that data are collected through cross-sectional approaches and that the ratios of effectively measured prevalences between exposed and unexposed are always 2.0. Bearing a causal outlook, the fundamental issue concerns the interpretation in terms of the experience of the population under investigation. In the light of some or even all of the conditions presented (c.f. subsection “Purpose of cross-sectional studies”), how are the estimates to be read? What do they signify in terms of CI and ID, and by extension, their related effect measures CIR and IDR?”

2. Minor Essential Revisions

2.1 I would delete the words bias and unbiased in the abstract, since I use these to refer to estimators, which are not discussed explicitly in the paper.

Done

2.2 English errors

- Page 2, line 5: settled are >> settled and
Corrected.

- Page 2, line 7: such the >> such as
Corrected.

- Page 3, line 2: methods >> methods are
We opted to keep the original since introducing “are” would alter our intention to convey that “Mainstream books ... used to emphasize the study of ...” and not ‘are used to’.

- Page 3, line 14: , whether >> :

We opted to keep the original since introducing a colon where indicated would repeat this punctuation in a single sentence. We hope this is acceptable.

- Page 3, line 16: in multivariate >> in the multivariate
Corrected.

- Page 14, line 10: in are >> in ??? are
Thanks for pointing this out. The passage now reads: “... to ask whether the structuring assumptions previously outlined are actually met.”

- Page 14 line 12: not least important >> not important
We opted for using the word ‘less’ since removing ‘least’ would convey the opposite message we intended, which is that a descriptive perspective is important from a public health perspective.

- Page 15, line 23: akin a density >> akin to a density
Corrected.

- Figure 4, line 5: extra bracket to remove.
Corrected.

3. Discretionary Revisions

None

Reviewer #2 (TB)

1. Major Compulsory Revisions

1.1 The authors’ main argument is that a cross-sectional study (CSS) should be considered as a kind of nested-case-control study (CCS) for which the POR would be the ‘best’ estimator of the incidence density ratio while the PR would overestimate risks in most cases. To demonstrate this, the authors present simulations with respect to several aspects: frequency of disease, equality of duration of disease among exposed and unexposed subjects, and duration of follow-up of the presumed underlying cohort. The cumulative incidence ratio (IR) only met the PR in a few restricted scenarios. Neither did the PR equal the incidence density ratio (IDR), which in contrast the POR did. In contrast to the authors’ notion, I think that the majority of CSS will meet the restricted scenarios and that a PR should not be interpreted as an IDR in these study settings. While I agree with the authors that the original purpose of a cross-sectional study (CSS) should be the study of prevalences, it has become increasingly common to use this study type for causal inferences. Cross-sectional studies in occupational and environmental epidemiology have gone so far to study associations between presumed risk factors and one single diseases entity (instead of many which the cross-sectional design would be most suited for). So the purpose of cross-sectional designs has shifted towards studying causal relationships, even if one is aware of all the associated weaknesses of this approach. The authors, legitimately opting for longitudinal designs, argue that studying the ratio of two prevalences may conflate the incidence and duration of the disease. While this
notion is undoubtedly true under the circumstances of a follow-up in time, we still have to
decide about the most natural and interpretable ‘risk measure’ for CSS which measures
disease and exposure in a certain point of time. CSS, in this respect, face similar problems as,
e.g. a case-control study, particularly if it is not nested into a cohort.

These general comments are addressed in the ensuing responses since they are taken up again in
what follows.

1.2 The authors provide five conditions to interpret CSS, of which three (stationarity, non-
selective survival between exposure groups, and same duration of disease among the studied
groups) can be assumed for most CSS, while the other two (temporal direction and reverse
causality) are inherent to any cross-sectional design and should therefore be of no interest in
deciding whether a PR or POR is more suitable as a measure of association.

While sustaining that all five assumptions need to be considered when scrutinizing whether a CSS
enables causal inference, we do agree that temporal direction and reverse causality do not directly
guide any decision as to the use PR or POR. We have thus opted to explicitly point this out in the
text.

1.3 A CSS is unable to assess the duration of disease. However, studying diseases with short and
changing outcomes (such as infectious diseases) does not make a lot of sense in a cross-
sectional setting. CSS commonly deal with chronic diseases which renders the question of
duration of disease not important. Accordingly, duration of follow-up is not captured by a
CSS, as it estimates an association between a disease and a potential risk factor at a given
point in time.

CSS’s are surely unable to assess the duration of disease. Still, the pattern of durations among
exposed and non-exposed is crucial to bear in so far it may or may not render a CSS capable of
producing a measure that is meaningful within a causal framework (such as the potential outcomes
model). Although disease duration across exposure strata are not measured directly, their equality
has to be assumed a priori —preferably stemming from a substantial background knowledge about
the outcome event under study— if one is to proceed with using cross-sectional data to address
causal (explanatory) relations. This is evident in equations 10 (page 9). Although durations are a
part of the equation, it is not necessary to know their values and to actually measure them in a CSS.
If the equality assumption is tenable, they can cancel out across groups and enable a meaningful
estimate (in terms of a causal incidence contrast) to be calculated.

1.4 Thus, by this definition a PR is rather similar to the cumulative incidence ratio (CIR), but not
to the incidence density ratio.

We fail to understand why “PR is rather similar to the cumulative incidence ratio (CIR)” given
“duration of follow-up is not captured by a CSS”. Is this because the reviewer is considering that
both PR and CIR compare “two disease frequencies at a certain point in time” rendering the
duration of follow up irrelevant? If so, allow us to disagree since all risk measures (cumulative
incidences) are, by definition (c.f. equation 1, page 8), always referred to a given follow up period
(whilst the PR does not since it relates to prevalences — state of disease at any given point in time).

Revisiting page 11 of the current (second) version of the paper, we write (abridged):

“[…] it is never possible to specify to which At-risk the PR really relates to. Having
carried out a survey and estimating the contrast between two prevalences (exposed and
unexposed), the researcher will always be in dark as to which time period the risks
account for and thus, by extension, the ensuing CIR purportedly emulated by the PR”.
Thus “[…] in a survey, what would the interpretation of a prevalence contrast be in
terms of a CIR (relative risk) on detecting, for instance, a PR = 2.0 as a proxy to a
CIR = 2.0 […]? […] From the stance of a relative risk (CIR) interpretation, a
researcher uncovering a PR = 2.0 cannot know which $\Delta t$ is at issue. The assumption that $\Delta t = 1.45$ [this is value used in the example] is empirically unrecognizable, as is thus the interpretation itself. Counter to a common view, a detected prevalence estimate may not be referred to any risk estimate. The key point is that they have little bearing to the CIR, which comes to show that [...] the interpretation of PR in terms of a CIR also implies a conceptual misunderstanding.”

1.5 Even if the PR is able to capture the CIR in a narrow range of scenarios only, those will be met by most CSS.

From the above argument (R#2, § 1.4), this passage does also not hold since, accordingly, the PR does not capture the CIR in most circumstances. An approximation of these two quantities is only achieved when the disease is rare, a situation that is quite unlikely to be studied with a CSS.

As to the scenarios being met by most CCS, if the reviewer is referring to those outlined in section “Structuring conditions” (new version), allow us to diverge once again. Regarding specific postulated exposure-disease relations, in many CCS (1) stationarity cannot be upheld; (2) selective survival is quite likely; (3) equal duration of disease across exposures (and the auspicious consequences in terms of a meaningful effect measure; § 1.4) is not sustainable; (4) the ‘no reverse causality assumption’ is hard to assume; and (5) an a priori assumption that the exposure is effectively an antecedent of the outcome may not be easy to maintain. As much as we believe that CSS are indeed useful in addressing causal questions, we also believe that researchers need to evaluate in which circumstances inferences are appropriate. Alas, many CSS preclude ‘causal modelling’, not because of hindrances at the analysis stage, but right at the outset when one or more of the structuring assumptions are violated.

1.6 However, interpreting the PR as a natural and comprehensible estimate of a CIR (comparing two disease frequencies at a certain point in time), the rare disease assumption becomes important again, and a POR would rather overestimate an association if interpreted as a relative risk.

If the central tenets of our paper are to be hold, this passage would not be consistent. For one, as we hope to have shown, the PR may not be “natural and comprehensive” estimate of the CIR since the latter is totally dependent on a given (specified) $\Delta$-risk (c.f. R#2, § 1.4 and § 1.5), whereas the PR, as naturally based on two prevalences, has no bearing on any given follow up period. It follows that since the POR is not intended to (and cannot) capture the CIR but rather the IDR in certain circumstances, the disease rarity is not an issue here. As sustained on page 15 of the text, this is because certain CSS’s may be viewed as a density sampling case-control study wherein non-prevalent cases are selected proportional to respective person-time quantities, and as well known for long, the disease rarity becomes unimportant here.

1.7 Researchers have to be aware that CSS suffer from many limitations incl. the lack to establish causality as well as the temporal inferences outlined in the article. If they are aware that CSS are not a particular suitable study type to analyze risks, adjusting for other factors may add to the explorative nature of these studies. Later on, the CSS may give rise to a more suitable design such as a longitudinal study.

We agree that, despite certain limitations regarding causal inference, the CSS may be “explorative in nature” and “may give rise to more suitable designs such as longitudinal studies”. Yet, engaging in controlling confounders only adds if rationalized and applied within a causal framework from the start. We hold that there must be a causal model from the outset (which incidentally does not depend on any study design), but once data from a CSS is at hand (to be used), there is a need to further specify the conditions —be it temporality or any of the others raised in the paper— whereby causal inference is sustainable. Thus, modelling data from CSS (incl. adjusting for confounders) should not be an unrestricted approach, but rather a procedure that may only be implemented once
certain conditions are scrutinizes and deemed to hold. In fact, this issue is also one of the essences of the paper.

2. Minor Essential Revisions

2.1 Abstract

- Replace ‘are’ by ‘and’ (1st paragraph, last sentence)
  Corrected.

- Add ‘such as’ (2nd paragraph)
  Corrected.

2.2 Background:

- Write prevalence ratio and prevalence odds ratio in small type (2nd paragraph)
  Corrected.

- p. 13: The calculation of a Cox regression was rather used as an adaptation for cross-sectional studies than based on the idea to consider the CSS as a retrospective cohort. In this approach the follow-up time is set to one in all subjects. (Under these circumstances Cox is also mathematically identical to a Poisson regression).

  We know this and agree. The passage only mentions that the Cox regression has been proposed and used.

3. Discretionary Revisions

List of abbreviations: Put in alphabetical order.

Done.