Reviewer’s report

Title: Novel use of three administrative datasets to establish a cohort for environmental health research

Version: 2 Date: 15 December 2014

Reviewer: Louisa Jorm

Reviewer’s report:

This is a carefully-crafted and well-written article that investigates the utility of a national health register as a total population denominator for epidemiological studies. Although its applicability is limited to the New Zealand context, this is exactly the sort of careful empirical work that we need to underpin the effective use of administrative data resources for population health research.

Minor essential revisions

The authors state (background, line 52) that “there has been little description of how such data have been used”. This ignores the *very* many published papers using linked administrative data (from Canada, Scotland, Australia and Scandinavian countries, at the very least) to answer a plethora of research questions. Maybe the deficit is actually in published work that explores the quality and validity of linked administrative data, rather than work that “uses” these data to answer important research questions? I suggest that this text be revised accordingly.

Much more detail needs to be provided about the procedures used for matching the datasets (e.g. probabilistic versus deterministic matching, software used, thresholds for clerical review, estimates of false positive and false negative matches) and, importantly, the implications of these methods for the generalisability of the findings. To make this paper important beyond the New Zealand context, this detail is critical!

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I declare that I have no competing interests