Reviewer's report

**Title:** Enhanced reporting of deaths among Aboriginal and Torres Strait Islander peoples using linked administrative health datasets

**Version:** 2  **Date:** 30 March 2012

**Reviewer:** Stephen Guthridge

**Reviewer's report:**

Thank you for the opportunity to review this paper. The paper proposes improved identification of Indigenous status in statistical reporting of deaths data by the use of linked health datasets. The difference in this study, is the authors’ use of an algorithm for deciding Indigenous status in multiple datasets, rather than accepting a single report of an individual being an Aboriginal of Torres Strait Islander in any one dataset, as has been proposed in previous similar linkage studies. This is an incremental refinement, which directly addresses the risk of overestimation of the Indigenous population, when it forms only be a few percent of the total population.

The methods used in this study are uncomplicated and appropriate. The data used in the study are routine administrative health datasets including deaths data.

The discussion is coherent and relates directly to the results. The discussion and conclusion are modest in their claim and do not seek to overstate the importance of this refinement. The study builds on and acknowledges the established work in this field.

Discretionary revisions

There are two additions which I propose will improve the paper.

1. The first is that the discussion is limited to the applicability in NSW. The paper would benefit from wider discussion on the relevance of this approach to other Australian states, in which Indigenous status in deaths data is either more or less complete than in NSW.

2. A second addition, or perhaps a correction to the text, is the impact of enhancement on the number of deaths in younger age groups. In paragraph 2 of Results, the authors report that “apart from those less than one year of age, there is little increase in the number of reported deaths among Aboriginal and Torres Strait Islander children and young people”. However in Table 1 the increase is reported as between 0 (age 5 to 9) and 100% (age 10-14), with many age groups up to 60-64 having increases of 25-30% using the algorithm and greater for “at least 1 report”. These changes are not inconsequential. Importantly, any differential increase, if it exists, associated with age will directly impact and perhaps bias life expectancy estimates. A differential enhancement of deaths in
the elderly over the young will have an even greater effect on the estimation of Median Age At Death. How much of the improvement in MAAD reported in Table 2 for example is a consequence of this (possible) differential enhancement?

The authors need to review their interpretation of these results. If there is a potential bias in the enhancement method then the significance should be acknowledged in the discussion.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.