Author's response to reviews

Title: Health behaviour modelling for prenatal diagnosis in Australia: A geodemographic framework for health service utilisation and policy development

Authors:

Evelyne E Muggli (evi.muggli@mcri.edu.au)
David McCloskey (davidm@pathfindersolutions.com.au)
Jane L Halliday (jane.halliday@mcri.edu.au)

Version: 4 Date: 7 July 2006

Author's response to reviews: see over
Health behaviour modelling for prenatal diagnosis in Australia:
A geodemographic framework for health service utilisation and policy development

Muggli, Evelyne Elvira, evi.muggli@mcri.edu.au,
McCloskey, David, davidm@pathfindersolutions.com.au, and
Halliday, Jane Lavinia, jane.halliday@mcri.edu.au.

Authors’ response to reviews

Dear Editors,

We would like to thank reviewer Babak Khoshnood for accepting the revised version of the paper. His expertise and comments have helped to improve the original manuscript greatly. We have amended Figure 3 as suggested by reviewer Joan Morris and we believe this has also added value to the paper.

Given that reviewer Jane McElroy raised so many concerns in her initial review and found additional concerns in her second review that she did not raise previously, it is quite possible that we will not be able to satisfy these even after a second revision. If this were the case, we would be disappointed in view of the assistance and support of both other reviewers.

However, we hope that the revised manuscript together with our response will clarify most of the concerns of this reviewer and assist the editors in making a decision.

Please find our response below, and the revised manuscript with highlighted changes attached for further consideration. We look forward to hearing from you soon.

Sincerely,

Evelyne Muggli
Public Health Genetics
We thank the reviewer again for their considered review of the above manuscript. Please find our response below.

**Minor essential revision**

1. **Figures 3 and 4 should have the live birth ratio of DS plotted against uptake of prenatal diagnosis to increase usefulness:** We thank the reviewer for this suggestion. We have now combined figures 3 and 4 and present the data as suggested in a revised Figure 3. This modification was followed by a number of minor changes to the text in the manuscript, which are highlighted in yellow.

2. **a) Formal statistical analysis:** Given that the data were modelled to the geodemographic segments we believe that our descriptive analysis was more appropriate. While it may be possible to apply statistical tests to our data, we wanted to avoid over interpretation of results derived from data of limited statistical validity (see discussion on street addresses).

   **b) Observations were several magnitudes higher than actual numbers:** Each Local Government area (LGA) is made up of a varying number of Census Collection Districts (CCDs), allocated to geodemographic segments based on census data and consumer surveys. A number of different CCDs from different LGAs may be combined in a geodemographic segment. Census collection districts (CCDs) are usually allocated using street addresses; however, these were not available in our data, so therefore a modelling approach was necessary, rather than proportional allocation of CCDs. For example, a single LGA(a) may consist of 3 CCDs (a1-a3), each of which may end up in a different geodemographic segments. The modelling approach allocated 100% of births and prenatal tests in the LGA(a) to CCD(a1), CCD(a2) and CCD(a3). The resulting numbers in the geodemographic segments were therefore multiples of the original numbers.

   We hope to have clarified the method of the modelling by altering the text in the manuscript as follows: **As this variable is not routinely collected by the PDCU, it was necessary to interpolate the LGA data and re-distribute this data to the component CCDs in each LGA. The modelling for this involved first taking the rate of testing that applied in each LGA and applying this standard rate to each component CCD.**
Manuscript Number 4576767529476271

Health behaviour modelling for prenatal diagnosis in Australia: A geodemographic framework for health service utilisation and policy development

Reviewer: Jane McElroy

We thank the reviewer for her review and comments on the above manuscript. Please find our response below. We have bulleted Jane McElroy’s comments (See Response, Attachment 1, pages 7-11) to facilitate our response and changes to the text are highlighted in blue.

General:

1. We believe that we have addressed and incorporated the original comments of this reviewer that led to an improvement in the manuscript. Given that the reviewer read the revised manuscript and commented on this before reading our response, we will now address her mostly new comments in the present response.

2. The aim of this study was to use large population-based, high quality data collections on births and prenatal diagnostic tests, with some well defined geographical indicators available, but without street address. It could not be a traditional sophisticated geographic information systems analysis without street address. We therefore think it inappropriate to calculate confidence intervals on modelled data that all have highly significant p values because of the magnitude of the adjusted sample size (see point 2b for previous reviewer). The presentation and discussion of the results were of a descriptive nature and interpretation must be within the limitations of the study. Nevertheless, the main outcome shows that there are specific areas in our State where there is differential uptake. In our conclusion we propose that this implies that there is a need to target health professionals and pregnant women in specific locations to ensure there is equity of access to services and that all pregnant women can make informed choices that are best for them. The other point we wanted to make with this paper was that this geodemographic framework is a potentially useful tool for policymakers and service providers and that we advocate for appropriate data collection. We believe we have not over interpreted our results.

3. We can only reiterate that we have presented a descriptive analysis. To clarify this point further we have added this phrase to the methods paragraph in the abstract. The revised Figures 1-3 visualise the results in a meaningful way and removing these would make the paper pointless.

Major Compulsory revision

4. Please see our response in point 2.
5. Please see our response in point 2.

6. We have added the following sentence: **This was done to ensure that the births and prenatal tests occurred within the same year to facilitate record linkage.**

7. We hope to have clarified this concern by amending the sentence as follows:
   Denominator data were obtained from all births (to calculate a Down syndrome live birth ratio) or confinements (to estimate uptake of prenatal diagnosis) recorded at the PDCU in 1998 and 2002.

8. The geocoding method and postcode assignment are outlined in the Methods section with more detailed reference to the commercial geocoding process presented as an additional file. See also Point 2. To emphasise the validity of the postcode assignment, we have added the following sentence: **The allocation of postcodes in Victoria underwent a major change in 1996 followed by the creation of new Local Government Areas (LGAs) which have remained stable since.**

   In regard to transcriptional errors, we have inserted a sentence in the description of the database. **Double-entry and other ongoing validation activities ensure data of the highest quality and reliability.** We have supplied a reference for our validation work which is of course essential and part of our routine quality control activities in Victoria.

9. We received our raw data, including LGAs, from the Perinatal Data Collection Unit of the Victorian State Government. For completeness of reporting, we simply stated that they have a method to assign postcodes to LGAs. If this sentence continues to cause confusion we are happy to remove it, and report that we used the data, including LGAs, as we received it. As already stated, (point 2 repeated), the geocoding method was designed to be based on CCDs, not LGAs. Without street addresses we were unable to reliably allocate CCDs to each record. The method we used instead is described in the second half of the section “geodemographic segments”.

10. See point 9.

11. See point 9. We believe that the manuscript contains all the information pertinent to the study and in the interest of being concise further detail would continue to increase the word count without adding value to the paper.

12. We have added a footnote to Table 1: **Proportional increase in household income, compared to lowest in Victoria (Single Parent Public Housing).** The grading of
household income in this way makes the magnitude of the increases between segments accessible to the reader.

13. Again, the word limit of this manuscript is in accordance with the publication in a scientific journal. We used commercially available geodemographic segments in our study where, based on their CCD composition, a number of segments are defined as rural, while others are metropolitan. We believe that they can be taken at face value.

14. The original numbers would be too small to do two-year maternal age intervals. The rationale for maternal age stratification is clearly described on page 4 in the Introduction and we feel we should not have to repeat this

15. See Point 2.

16. We used the maternal age of 37 years as a cut off in our study because this is the age at which prenatal diagnostic testing is offered (based on advanced maternal age alone) in Victoria. The sentence referring to a 35-year maternal age cut-off is of interest to international readers in the field as this is a common age cut-off for advanced maternal age in other countries. We therefore believe this sentence should not be deleted and have inserted the following qualification: When the advanced maternal age cut off was set at 35 years as it is done in a number of other countries.

17. Tables 2 and 3 are common baseline descriptions of the data.

18. This is described in the second half of the study population section.

19. Births and prenatal diagnostic tests had to be linked by probabilistic record linkage. We selected those years as described above and in the text of the manuscript. The data on births of babies born with Down syndrome were pooled from 3-year data to obtain a larger sample size and account for annual fluctuations routinely seen in Birth defect data.

20. See point 2.

21. See point 2.

22. There are two centres that offer prenatal diagnosis in country Victoria and women can travel up to five hours to a metropolitan or rural centre for testing. First trimester combined screening is not available to most women in rural Victoria but we have no figures on the actual proportion. We believe that for the purposes of the introduction, women have “few options but to travel to a metropolitan centre” is sufficient. We discuss access and opportunity in more detail when looking at the different segments.

23. Reporting the p values in the text of the manuscript makes it easy for readers to interpret the data as they read.
24. We have now labelled the Y-axes of Figures 1-3.

25. We have considered this point carefully and agree that education is not the only need, but is perhaps one of the most obvious. We have amended the sentence to focus more on equity of access and informed choice.
Bulleted comments by Jane McElroy, as addressed by authors in second revision

1. I am puzzled as to why some of the concerns or questions I had were not addressed in the text but rather provided in a note from the authors. I understand that some need not be part of the paper since they reflect my lack of familiarity with Australian geographic units and terminology and that might be somewhat unique. But others were appropriate in the revision.

2. â€œWe have decided not to offer any form of statistical analysis for the comparison of data relating to the geodemographic segmentsâ€‌ makes the paper considerably weaker and the conclusion even less strong than I originally thought. There are statistical methods to calculate CI on these data so I am puzzled as to why the authors choose to not do so.

3. If the authors choose to not test these data in any rigorous way, then I think providing graphs that show apparent but maybe not real differences between the segments may be misrepresentative and should be eliminated to lessen misunderstanding about the results. As stated in the section responding to the abstract: â€œA lack of statewide geodemographic consistency in uptake of prenatal diagnosisâ€‌ can stand if we very very loosely define consistency to mean more or less exactly equal values. Of course, this phrase implies there is a meaningful difference in the different segments which absolutely is not supported in these analyses. Whether there are substantive differences or not remains to be explored.

4. Abstract: The conclusion speaks beyond the analysis. Making a concluding statement about the need for appropriate health service provision for families of children with Down Syndrome is fine but it was not what the paper was aboutâ€”uptake of prenatal diagnosis and prevalence of Downs Syndrome. The paper did not analyze data regarding the distribution of appropriate health service provisions for families with Down Syndrome.

5. The overall conclusion: â€œA lack of statewide geodemographic consistency in uptake of prenatal diagnosisâ€‌ can stand if we very very loosely define consistency to mean more or less exactly equal values. Of course, this phrase implies there is a meaningful difference in the different segments which absolutely is not
supported in these analyses. Whether there are substantive differences or not remains to be explored.

6. All women residing in Victoria who had an amnio or CVS in 1998 or 2002 and who expected date of delivery was also in that year were included in the study. Please respond to the concern about inclusion criteria in the paper—it is not sufficient to read the response to the concern in a note back to the reviewers. I suspect that if I have this confusion, there will be others.

7. I struggle with the terminology. Denominator data were obtained from all recorded births (Down syndrome prevalence) or confinements (prenatal diagnosis) at the PDCU in 1998 and 2002.

Do the authors mean to say that the denominator in their calculations where they use this denominator data is DS prevalence—(the number of DS births in 1998/number of person (child bearing women?) in population in 1998) x 100. I am still not clear what the denominator is and it seems very odd that the denominator would contain a prevalence calculation so I think it may be just syntax issue that needs to be fixed.

8. I find it unacceptable that the geocoding method, particularly the match rates and sensitivity scores that provided a match was not described. The entire project rests on the premise that the geocodes were accurate and that the women/births were correctly assigned to the proper LGA. Without this information, the quality of the analysis cannot be fairly determined. For the US, postcodes change over time and I wonder if this is an issue in Australia and if so how this was dealt with in the geocodes. Further, it seems implausible that all the postcodes contained zero errors? How were obvious typos dealt with.

9. And where a number of postcodes are divided between two or more different LGSs, allocation to a specific LGA was assigned proportionally by the PDCU on the basis of Australian Bureau of Statistics census data. This sentence lack precise description of what was done and needs to be clear to the reader. In the authors response, they note that they respond to this concern by saying a limitation is lack of street addresses. Of course that is a clear limitation; however, the authors analyzed the data and assigned geocodes somehow. The method that was used to do this needs to be explained.

10. If I took a stab at figuring out what the authors did, based on my experience with this problem, then I would say, the sum of number of births by postcode was determined from birth registry data for years 1998 and 2002. For postcodes in which their boundaries bisected one or more LGA, the number of births was allocated to each intersecting LGA based on ??????????
It could be based on the percent of area that bisected the LGA or it could be based on something like what US census uses—proportion of block groups within each bisecting boundary, or it could be based on the proportion of each postcode population or â€¢.

11. After reading the authors response to my concerns, it seems the authors divide the birth data, which is summed by postcode, as a proportion of the total population of the LGA? Correct? If so, this needs to be stated and it would be better to divide the data based on population of child bearing aged women. If postcode 3351 had 100 births, then Ararat LGA = 8 births, Ballarat LGA= 14 births, Golden Plains LGA = 61 births, and Pyrenees = 16 births from this postcode. Please describe what was done in the manuscript.

12. Please describe the calculation to obtain the numbers in Table 1 for â€œgraded household income.â€œ I would guess that the denominator is the income for the lowest category. If so what is that? What if the difference between 1.49 and 1.51? How much difference in $; is it $1000 or 10,000? With no data, the magnitude of the difference between the categories and between high medium and low categories cannot be determined. Please provide an example and if my guess was right the denominator $ amount used in the calculation.

13. Please also describe how the â€œBirths (% of metro or rural) was calculated and what it means in the text. Again what is the denominator for these percentages, total number of births in the metro area? Again, another definition is needed. How was rural and metro defined? By LGA or postcode at xx population or xxl child bearing female population or what?

14. For modeled rates of prenatal diagnosis, was this calculated by age or by the two age groups: under 37 and 37+ years. Please specify in the text. Iâ€™m guessing that it was just calculated by the 2 age groups and if that is so a different answer might occur depending on the age structure of the segment. I would encourage the modeled rates to be based on age by 2-year intervals (as was done in Table 2) since the risk dramatically increases with each year. I donâ€™t know the distribution of the mothers ages by segment, so I donâ€™t have any idea how stable the numbers would be by segment if the data were calculated by 2 year groups. If the clumping is due to small numbers this needs to be noted in the manuscript.

15. Another important component of graphing these rates in providing confidence intervals. It is unclear whether there is any real difference between any of these valuesâ€”all depends on CI. Even within the urban or rural segments, it is unknown whether there is any real difference in any of the values. This picture without CI tells
one story” that urban and rural are really different and there is limited variability within the two stratum--but I would not be surprised if very few if any of the segments were different from each other and that the strata are similar.

16. Since the whole analysis revolves around the cutoff at maternal age of 37, it is not clear to me why the authors report an alternative cutoff point to report results of change in percent of maternal age except possibly due to a different cutoff point being statistically significant whereas the one used in the analysis, age 37 was not. Delete this sentence. “When the advanced maternal age cut off was set at 35 years, increasing significantly to 46.7% in 2002 (p=0.04, not shown).

17. Table 2 doesn’t match the methods as far as I can tell.

18. “All women residing in Victoria who had an amnio or CVS in 1998 or 2002 and who expected date of delivery was also in that year were included in the study. This was in the methods section describing the study population. Where does 1997, 1999, 2000, 2001 come from? OK to include it but it needs to be described properly in the methods section.

19. Secondly, I don’t understand why the authors are reporting different time periods for births vs DS births in Table 2. Why do they not report the same periods, particularly for all births.

20. The same comment about the need for CI for figure 1 and 2 also holds for figure 3 and 4. Once the CI are taken into consideration, then reporting the results might substantively change.

21. Since the CI of these analyses need to be taken into consideration before conclusions can be drawn, I cannot comment in any meaningful way to most of the discussion.

22. 2nd paragraph, last sentence: “while second trimester maternal serum screening is available throughout Victoria, all pregnant women residing in rural regions have few options but to travel to a metropolitan centre to have a prenatal diagnostic test” I What does “few options” mean more precisely? Can rural women have a prenatal diagnostic test somewhere beside metropolitan centre? How far do rural women have to travel (average)? Â½ hour to a location, 2 hours to a location? Although in the next paragraph, the authors somewhat dismiss this characteristic“travel time, please be more precise in characterizing the situation since an attempt was made to describe travel time.

23. In the text, I would not bother to report p-values for the table 2 and 3 results. I don’t feel it contributes in a substantive way and having it in the table is sufficient.
24. The y-axis on all figures need to be labeled (if these figures are kept)

25. Second point in the conclusion from the abstract and manuscript that the authors may or may not agree with is in my opinion the tiresome call for more education as an action plan. In fact, implicit in the push for education is the assumption that lack of knowledge (or ignorance) is what is driving people’s decisions. As highly educated people, it is hard not to think this. "If they only knew!" However, I would argue that many variables comprise decision making and it isn't as simple as educating people to affect behaviour change.