Reviewer’s report

Title: Calculating census tract-based life expectancy in New York state: a generalizable approach

Version: 0 Date: 21 May 2017

Reviewer: Robert Knorr

Reviewer’s report:

As an advocate of the value in estimating Life Expectancy at the sub-county level and as a public health practitioner familiar with the challenges for doing so, I found the paper by Tom Talbot and his colleagues to be an excellent real-world guide to how to face the technical challenges of estimating life expectancy at the sub-county level and for successfully overcoming them. The description of innovative methods developed at the New York State Department of Health by the author was clear and graphics enhanced the narrative greatly. The paper, with some minor revisions, should serve as a useful tool to others who have an interest in better understanding the disparities in health we have observed across the country.

Main Comments

1. The paper presents four challenges in estimating life expectancy (LE) at the census tract (CT) level (lack of accurate geocoding at the CT level, inconsistencies in where health departments record deaths versus where the US Census enumerates the populations used in the calculation of LE, a lack of access to death certificate data for residents who died out-of-state, and sparsely populated CTs have LEs with large margins of error) (beginning line 41, page 4). The authors state that how these were overcome would be addressed in the paper. While the author described each of these issues, the methods to overcome them were only presented for two; how to improve the accuracy of geocoding and how to overcome the sparse population limitation. So as to not limit the value of the the described strategies for overcoming these two important issues, I suggest that the stated purpose of the paper emphasize just these two issues. The other two can be presented just as they are elsewhere in the paper but as associated problems whose solutions were mostly unique to New York.

2. There is some inconsistency in the Abstract and Background sections. The Background section describes the goals mentioned in my comment #1. However, the Abstract states the paper provides a stepwise approach for calculating LE. I believe that the Abstract's purpose most clearly presents what the paper is about and suggest that this statement be added to the purpose stated in the Background (along with the other suggested edits mentioned above) and that mention of the geocoding and small numbers methods be used to enhance the Abstract.

3. The paper's value is clearly as a methods paper. I felt that the presentation of LE results as findings in and of themselves rather than in the context of demonstrating methodological concepts to be a distraction in the paper, especially since the scientific rationale for the analysis
presented were not be adequately introduced. A reference on the basis for looking at the selected variables would be sufficient, along with the presentation of the results as illustrations of the application of the methods being presented (as the results were addressed in the Discussion). I would have liked to have seen, instead, a table contrasting results seen when no solutions to the challenges are applied and/or whether findings would have been different if LEs were only presented at the county level.

4. The Abstract mentions the rationale for looking at LE at the CT level, but this was not followed up on in the full paper. Only brief mention is made in the Background section and none in the Discussion or Conclusion. It is recommended that why a reader should care about estimating LE at the CT level and it's added value over existing LEs at the county level should be presented in the body of the paper.

5. The last line of the Abstract concluded that "Reliable fine geographic scale estimates are needed to aid public health officials better focus preventative health programs in high risk populations in both rural and urban populations". However, no data were presented to demonstrate that rural populations or high risk populations were preserved through the aggregation of tracts to allow the LEs to aid in program planning for them. A real-world example would help support this conclusion.

Other Comments

1. The paragraph beginning on line 12 on page 5 would be better placed in the methods (and discussion) section. This is because the method described is only a potential solution to one of the stated challenges. I suggest especially moving it to pg 7, line 56 of the Discussion.

2. The significance of selecting a Standard Error greater than 2 might benefit the reader.

3. The Abstract (last line) refers to a value of LE estimates in planning public health programs in rural and urban areas. Neither the Discussion nor Conclusion returns to this point. Further discussion to support this point would benefit both the application of the Geographic Aggregation Tool and LE estimates at the CT level.

Level of interest
Please indicate how interesting you found the manuscript:

An article of importance in its field

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:
1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare a potential limited competing interest based on criterion 6. The author and myself both served on an CDC Working Group, the Sub-County Assessment of Life Expectancy (SCALE) Project and one of my review comments concerns a perceived lack of acknowledgement of that groups impact on the work

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal