Reviewer’s report

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

Version: 0 Date: 28 Jun 2017

Reviewer: Christopher Paulu

Reviewer’s report:

LE in NY - Pop Health Metrics – review

Overall

This is an engaging and useful demonstration of obtaining an important population health indicator (life expectancy) at the level of census tract, which is more relevant and actionable than county or state estimates.

There are three areas of methodology that require more justification, discussion, and/or detail:

1 - aggregating regions based on numerator/deaths is akin to "Texas sharpshooting." It would tend to systematically bias mortality rates upward. What is the impact in this exercise, and how is the approach justified? What are the alternatives?

2 - excluding tracts with populations mostly living in group quarters deals with the systematic bias applicable to those tracts. However, it cannot address the deaths among the group quarters' populations that were coded elsewhere, which would give an upward bias (elsewhere). Please discuss the nature and extent of this bias. One useful metric would be to include an LE estimate for upstate NY in total, compared to an overall LE estimate for the final study population after exclusions. Further questions: Would it be feasible to re-assign deaths to the place of residence per census definition? Or conversely, assign group quarters persons to where their death would be assigned?

3 - a second stage aggregation (or aggregations?) was performed until all units had an SE <2 yrs. How were these combined? Were existing units with SE <2 excluded? Please provide more descriptive detail so that others can replicate. It's difficult to evaluate this step of the approach without further details.

General comments/questions

[p. 2, Line 11, and elsewhere] The language around geographic scale and population size is loose, tends to equate these aspects, and muddies the intentions. Nations, states, and counties are geopolitical units, not geographic scales (e.g., Rhode Island is not the same scale as Montana). At the same time, units of similar geographic scale can have vastly different population scale.
What is the purpose of the present exercise? Why change from geopolitical units to census sub-units? Is it to assess small areas, or small populations? Or neither, really: this reader presumes the intent is to approximate a unit closer to "community" level, and that community-level public health indicators provide more precision for action, as well as uncovering disparities otherwise masked.

It would be useful to answer the question "why did we use census tracts?" They are clearly useful in providing associated demographic variables. But how well do census tracts serve the purpose of approximating communities? Some discussion on how the US Census defines and constructs tracts could be useful in the Background or Methods sections.

[p.3, 27-30] "Life expectancy...summarizes mortality patterns across all age groups." Please provide an alternative definition for the indicator. The public would understandably apply this measure to the future or an individual born in the population, yet the measure is not based on projections. So to whom does the years life expectancy apply?

[p.6] The authors seem to have invested a great deal of effort to geo-match a final 1.7% of addresses for the death records. It would be useful for others to know how many person-hours that effort took - and how much of a difference it made. Please address these issues in the Results/Discussion.

[p.7, 16-22] Please explain how the NY death registry defines or assigns residence "address." Would it be common to all death registries? Compare to Census definition. It would also be useful to have some examples of group quarters (or Census definition) early in this paragraph.

[p.7, 46] Please explain why having an SE <2 is important. Should an LE of 65 with an SE of 3 be rejected in the same manner as an LE of 80 with an SE of 3? In general, what "margin of error" for LE is acceptable, and why? For other measures, it's common to see RSE used as a criterion for stability; is there any similar metric available for LE?

[p.8, 7] Why were tracts merged within town boundaries (of a certain size)? Presumably this relates to the "community" relevance objective.

[p.8, 19-27] Please provide detail on how the final aggregation step was performed.

[p.8, 38-43] How would the ability to assess sociodemographic associations be impacted if the units were e.g. county? It seems worth mentioning that tracts break the population into groups with much wider variation in the sociodemographic variables of interest (or put another way, county units would combine and mask that variation).

[p.9, 37-54] Table 2: this table should probably exclude the column for percent of deaths. There are deaths to study population residents, such as those in NYC, that cannot be accounted for in Table 2.
[p.10, 19-31] The description of using the GAT seems to omit that the target was set to 60+ deaths. It would be useful to know, of the 2415 post-aggregated regions, how many remained individual tracts.

[p.11, 4-9] How was the # of deaths in NYC to NYS residents determined?

[p.11, 14] It would be useful to include the US LE for comparable time period here.

[p.11, 26-36] It would be useful to explain that (and how) this kind of surveillance analysis is enabled by having tract-based units (rather than e.g., counties).

[p.12, 7-24] Table 3: all the row/column labels need to specify the inclusive/exclusive boundaries, e.g. if "0-2%" is actually "0-<2%". The values would benefit from confidence intervals; it's not clear how stable the LE estimates are.

[p.12, 35] "no loss of subjects" - this sounds exaggerated given the exclusion of tracts and the loss of NYC deaths.

[p.12, 37] "Census tracts were found to be a useful building block…" - but why better or worse than block groups, or zip codes?

[p.12, 54-59] Populations were excluded, but their deaths, coded elsewhere, could not be. The number of those deaths is unknown.

[p.13, 12-15] What was SEPHO's reason for a threshold of 5,000 person-years? Was it an SE of 2? Did the SEPHO report indicate 5,000 persons for annual LE estimates? Did that report address the impact of aggregating years? Also, add reference citation to this sentence.

[p.13, 27-32] The utility and applicability of the given estimate (population required for 95% of tracts to have SE <2) is not clear. The estimate of the percent of tracts having SEs <2 is contingent on the actual distribution of population by tract. Suggest leaving this statement out.

[p.13, 34-42] Please comment on how or if the results of this approach in NYS would be similar elsewhere in the US. Should other states expect similar results, since tracts are somewhat population-standardized? Or should rural states expect more tract aggregations?

[p.14, 5-9] This is another opportunity to speak to the advantage of using census tracts to assess sociodemographics, and look at associations with LE.

[p.14, 14-17] Why does migration pose an issue of bias? No population is entirely immovable, so this would seem to undercut LE as a useful indicator. Is it not inherent to LE that its based on the age-specific subpopulations that happen to live in a place? This consideration of how migration influences LE can inform how we generally interpret the indicator: i.e., it's a snapshot assessment of a population with a place attribute, not a characteristic of the place per se. It does seem useful to distinguish between "a healthy place" vs. "a place where healthy people go" (or the opposites) - and note that LE cannot on its own help us distinguish the two.
"The SEPHO LE calculator only estimates the error assuming both the population and death counts are correct." This sounds like a situation with zero error. Does the SE assume a sampling distribution for death events? Also population? Please clarify.

"Our study area was large with over 2,600 census tracts." To clarify: the study area had a large number of base units - over 2,600 tracts.

The population threshold of 3,000 doesn't match the discussed thresholds of 1,000 population for 5 years, and a county of 3,000 persons might not be broken into multiple tracts. Would it be more informative to assess how many counties do or don't have census tract sub-units?

An important limitation of modeled estimates based on demographics is that these estimates become insensitive to changes in actual mortality. That makes them inappropriate for evaluating interventions. See 500 Cities project for discussion/acknowledgement of this limitation.

Conclusion. The abstract states that fine scale LE estimates "are needed to aid public health officials better focus preventative health programs in high risk populations in both rural and urban areas." This utility is not echoed or explored in the Conclusion section. The allusion to rural and urban is not explained. Divisions within metropolitan areas have been relatively well studied, while rural areas are generally assessed only by county?

Title: "Methods and Tools for Calculating Life Expectancy for Small Areas in New York State"

The title seems off the mark: the paper is not a guide to a large array of tools+methods, but rather a successfully applied approach with specific tools, that can be generalized to other e.g. states. Also, census tracts are not equivalent to "small areas," so perhaps the title should specify tracts as the focus.


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 that I have no competing interests

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