Author's response to reviews

Title: The effects of spatial population dataset choice on population at risk of disease estimates

Authors:

Andrew J Tatem (andy.tatem@gmail.com)
Nicholas Campiz (nicholas.campiz@gmail.com)
Peter W Gething (peter.gething@zoo.ox.ac.uk)
Robert W Snow (rsnow@nairobi.kemri-wellcome.org)
Catherine Linard (catherine.linard@zoo.ox.ac.uk)

Version: 2 Date: 20 January 2011

Author's response to reviews: see over
Dear Sir/Madam,

Please find enclosed our manuscript 'The effects of spatial population dataset choice on population at risk of disease estimates', which we have now revised in response to the comments of the referees and shown changes tracked. Below are our point-by-point responses to the comments, concerns and suggestions by the referees:

**Reviewer: Greg Yetman**

**Discretionary Revisions**

1. The review of GPW and GRUMP does not mention the availability of a 'mean geographic input unit' surface to aid in modeling uncertainty. See: http://sedac.ciesin.columbia.edu/gpw/docs/gpw3_documentation_final.pdf for details. Mentioning the availability of these data would underscore the importance of the statement on p12 that "...spatial population datasets should thus focus on integrating such uncertainties into the methods used for their construction as a priority".

   We have now included mention of this additional surface on page 12.

2. Maps could be clearer. Including a legend and units for all maps would be useful.

   We have now ensured that each map contains a legend and that the units are described.

**Reviewer: Stephen Lim**

*This paper describes an assessment of numbers that are paid little attention in the field of health measurement; that is, population estimates at small area levels. The authors demonstrate the importance of accurate population numbers by describing variation in the estimated population at risk of malaria that results from using different population datasets. This is an important topic that is line with...*
the objectives of the PHM journal. In general the paper is well written and the methods are sound and I have only limited comments.

1. **It is important to state in the abstract what spatial resolution the variation in population numbers is being assessed at.** I understood from the methods section that this is at the 5x5km level rather than the 1x1km resolution that some of the population datasets are available at (Table 1). Given that the finer the resolution the more potential there is for variation, it should be clearly stated that the assessment is done at the same resolution for all four datasets and what that resolution is.

We have now included in the abstract the spatial resolutions of each population dataset used for clarity. The reviewer is correct that finer resolution means more potential for variation, but the aim here was to explore the variations achievable using each dataset in its native form - i.e. as all those past studies in Table 2 have done. We could have (i) resampled the 5km surfaces to 1km, but this would have not changed results (unless some smoothing interpolation was used, which is not recommended for such datasets), or (ii) degraded the 1km surfaces to 5km, but this would have not replicated the findings possible through using the two most widely-used datasets, LandScan and GRUMP, nor provided a fair test of their capabilities for estimating PAR.

2. **In the assessment of population variation, the paper takes estimates of P. falciparum endemicity from 2007 and overlay this on population numbers from the four different sources.** For Landscan, these numbers are available directly, but for the other three datasets, numbers are projected to 2007 using intercensal growth rates. Are the Landscan numbers for 2007 (and the GPW3 numbers up to 2005) derived using the same intercensal growth rate method? How different would the comparison be if one were to take a year for which population data is available for all four variants (e.g. 2000) and applied the P. falciparum endemicity numbers?

The reviewer makes an interesting point here, however, all datasets generally use or are derived from similar growth rates, and thus the effects from any small variations in these will be obscured by the much larger variations in input data and spatial modelling approaches used. Moreover, as stated above and in the text, we are here aiming to replicate previous (and probably future) approaches to using these datasets, which have relied on intercensal growth rates to match population data to the year of epidemiological data. Finally, the main document focuses principally on comparing the variations in PAR derived from datasets for which the national totals have been adjusted to be identical, so as to examine the effects of the differing spatial modeling approaches.

3. **In the comparison of national-level assessments of PAR using the detail census data, the census data were available for different years (Mali – 2009, Namibia – 2001, Tanzania – 2002).** It was not clear to me how the authors resolved the temporal differences in this data compared to the population and PAR data that was for 2007?
This comment is a little confusing to us, since the 4th paragraph on P8 opens with 'For each country, the detailed population data were projected forward to 2007 to match the malaria data, using the same growth rates as described in the previous section.' This sentence seems to answer the referee’s question.

4. It would be informative to show the direction of the difference in Figure 4; i.e. +2.5% vs -2.5%. This would allow readers to assess whether CIA is systematically lower in terms of population numbers compared to UNPD.

This is a useful suggestion, and we have changed the figure to show the direction of difference.

5. Table S2 with an additional line for the global numbers could replace Figure 5.

We have replaced figure 5 with a version of table S2 that contains global numbers and only Landscan and GRUMP to maintain the consistency and focus of this section on these two most widely-used datasets. Table S2 containing comparisons for all datasets remains in the Additional File.

6. I found Figure 6 to be very hard to read – the size is very small given that multiple graphs/maps have been combined into one. I think a more intuitive way to present these numbers would be as a selected series scatterplots of the PARs>X% using the different datasets, e.g. GRUMP vs Landscan. Points could be color coded for the different regions (Africa, CSE Asia, Americas).

The scatterplots suggested by the referee are something that we have tried previously with presenting these results. These would work well if there weren’t a huge range in PAR numbers between countries - e.g. comparing PAR 5-40% for Nigeria vs Rwanda. One 'solution' to this is to use log-log scales on the scatterplots, however, this simply serves to obscure many of the key differences in PAR estimates between GRUMP and Landscan, particularly for countries with large PARs. After testing a variety of approaches to graphically representing the data and surveying a range of colleagues on their choices of graphics that got across the variations in PARs most clearly, the current graphs proved most popular - therefore we have elected to keep them in the resubmission rather than retread old ground.

7. Summary statistics are needed on the variation in PAR>X% at the country level between the four population datasets, e.g. a table presenting concordance correlation coefficients of the PAR>X% for Landscan vs Grump, GPW3 vs Landscan etc. This could also be done at a final geographical resolution, e.g. province

This is an excellent suggestion - we have calculated concordance correlation coefficients between the country-level PAR estimates of the four population datasets for each Pf transmission class and created
new tables for them. While these are useful statistics for readers to see, they do not add a huge amount
to the structure of the arguments in the main paper, so have been added to the Additional File. Whilst
the addition of a similar analysis at provincial level would be of interest, this goes beyond the scope and
aims of the current paper in demonstrating how national level summaries can vary through population
dataset choice. It would also add another table and more text to an already long paper. We have
however added concordance correlation coefficients to (what is now) table 4 of the main manuscript, to
provide an additional comparison metric for the detailed country-level analyses.

8. One of the implications that is mentioned is the need to “gather datasets into a central resource”. This
speaks to the need for increasing public availability of detailed census data. The authors make reference
to IPUMS but it would worthwhile making reference to current debates about increasing public
accessibility to health data, e.g. such as those covered in PLoS Medicine, the Lancet.

This is a good point and we have added text and references referring to these debates in the final
paragraph of the discussion.

We hope you find the revisions and responses acceptable, and we look forward to hearing from you
soon,

Best wishes,

Dr Andrew Tatem

Department of Geography and Emerging Pathogens Institute

University of Florida