Author’s response to reviews

Title: Characterizing measles transmission in India: a dynamic modeling study using verbal autopsy data

Authors:
Stephane Verguet (verguet@hsph.harvard.edu)
Ed Jones (edward.jones@lshtm.ac.uk)
Mira Johri (mira.johri@umontreal.ca)
Shaun Morris (shaun.morris@sickkids.ca)
Wilson Suraweera (Suraweera@smh.ca)
Cindy Gauvreau (Cindy.Gauvreau@partnershipagainstcancer.ca)
Prabhat Jha (JhaP@smh.ca)
Mark Jit (Mark.Jit@lshtm.ac.uk)

Version: 2 Date: 20 May 2017

Author’s response to reviews:

To: BMC Medicine Editors

Subject: “Characterizing measles transmission in India: a dynamic modeling study using verbal autopsy data” by Stéphane Verguet, Edward O. Jones, Mira Johri, Shaun K. Morris, Wilson Suraweera, Cindy L. Gauvreau, Prabhat Jha, and Mark Jit

Dear Editors,

Thank you for the opportunity to submit a revised version of our manuscript entitled “Characterizing measles transmission in India: a dynamic modeling study using verbal autopsy data.”

In detail below we give answers to the comments and suggestions the referees raised in their review.
Thank you very much for your consideration of our manuscript.

Sincerely,

Stéphane Verguet

Stéphane Verguet, MS, MPP, PhD
Assistant Professor of Global Health
Harvard T.H. Chan School of Public Health

Reviewer #1:

In general, I am happy with the changes made by the authors and would be happy to see this manuscript recommended for publication pending clarification of the simulation experiment done to test the performance of the pooled coherence procedure to estimate $R_0$ and $a_0$. I've given a decision of "major revision" only because it's not immediately apparent how to judge the simulation experiment and I'd like to see a clarification of this before rendering a final decision. The authors state that they simulated a 300 year time series, broke it up into 100 3-year chunks, and used each to estimate $R_0$ and $a_0$. This sounds valid, though I am naturally skeptical of the tight confidence intervals on the resulting estimates. Were these simulations from a deterministic model or a stochastic model? If the latter, how was noise added to the model? Was the model fit to the situated true incidence, or the simulated mortality; e.g. a binomial sample from the simulated incidence at a rate equal the CFR? What do these resulting time series look like relative to observed time series? The authors need to clarify these questions so that the reader can assess whether this is a fair comparison of performance -- e.g. if these were output from a deterministic model, of just the incidence of cases, then I might expect the method to perform really well -- but that would be an unfair test as the time series would have too strong a signal relative to the various sources of noise that appear in the real data.
Authors’ response: To address the reviewer’s concerns, we have expanded the description of the simulation experiment below, as well as expanding the scope of the simulations we conducted.

The 300 3-year time series were created by adding stochasticity to the deterministic model, in order to capture noise due to errors in the observation/measurement process.

The detailed process was as follows:

(i) Take the time series of number of cases each week over the 3 years from the deterministic model.

(ii) Simulate a time series of number of deaths each week by sampling from the distribution Bin(n,p) where p=1% (the assumed case-fatality risk) and n=number of cases in that week.

(iii) Add further noise to this time series of deaths by adding to each week a random number sampled from a uniform distribution U(0,m), with the value of m estimated below.

To estimate the parameter m in the uniform distribution U(0,m) we estimated the magnitude of noise we were seeing in the Indian (Million Death Study, MDS) data we fit the original model to:

(i) Generate a time series of measles cases from our SIR model using values of R0 and Amp with the highest coherence to MDS data (as described in our manuscript).

(ii) Convert this to the number of deaths by applying a case-fatality risk of 1%.
(iii) Scale the number of deaths in this time series to match an “MDS data set” by multiplying it by the ratio of the deaths in the MDS sample with the number predicted for the whole of India in Morris et al. (2013) [1]. We used the Morris et al. paper because this study used exactly the same MDS time series to estimate the number of measles cases in India during that time period.

(iv) Calculate the residuals, i.e. the difference between our scaled time series of deaths with the actual number of deaths each week in the MDS time series.

Using this process, we calculated that the average residual was 3,531 with a range of 0-19,897. Hence we sampled noise from a uniform distribution U(0,5000) which is roughly equivalent. Just to test our method to the limits, we also generated a time series with noise sampled from U(0,20000) instead.

The results are shown in the supplementary webappendix Figure S11.

Note that the raw MDS data were filtered to detrend the data using a Baxter King filter (bkfilter in R). This process has the effect of removing much of the noise at frequencies with small periods (<1 month) which are highly unlikely to be relevant to measles transmission, but does not affect the periods we were interested in (> 6 months). We apologize as we had inadvertently left this out of the description of our methods in the manuscript previously, and have now added it, as well as updated Figure 1 in the manuscript to show the effect of filter processing on the data and subsequent spectra.

With the noise term sampled from U(0,5000), we obtain R0 of 17.0 (95% interval 14.0-18.1) and Amp of 0.26 (0.12-0.41). With the noise term sampled from U(0,20000), we obtain R0 of 17.04 (13.4-19.5) and Amp of 0.25 (0.04-0.46) (Figure S12). This compares with the original R0 of 16.7 (95% Interval: 14.6-18.6) and Amp of 0.27 (95% Interval 0.14-0.41). Hence we find that adding uniform noise increases the uncertainty intervals around the estimates of R0 and Amp but does not have much effect on the central estimate. We show in Figure S13, that with the addition of uniform noise, there is little effect on the underlying harmonics.
Clearly there are ways to add noise so that it becomes difficult to estimate the underlying periodicity of the data. We could add uniform noise of such great magnitude that the underlying periodicity (the signal) is swamped by the noise; it would then be difficult to pick up any meaningful periodicity at all. We could add noise with its own periodicity or a truncated pulse of noise, such as seasonal reporting of cases/deaths, or data sampling of different efforts at varying periods. The spectral analysis of these time series would then pick up subharmonics in addition to the underlying periodicity. This is not surprising and would merely show that the method is working as intended, i.e. picking up periodic series without making any judgment about their origin.

As this paper was drawing from the MDS which has a high quality of reporting, we believe that MDS measles mortality data present less reporting biases and thus smaller noise, which enables the use of our method.

We have now clarified this in the supplementary webappendix (pages 30-32 and Figures S11, S12, and S13).

Figure S11. One of the 300 three-year sample time series of estimated deaths with added noise from either U(0,5000) (Figure S11A) or U(0,20000) (Figure S11B), respectively. Red line represents the data after Baxter-King filter.

Figure S12. The distribution of estimates of R0 and Amp, in left and right columns respectively, with added noise from either U(0,5000) (Figure S12A) or U(0,20000) (Figure S12B), respectively.

Figure S13. Longest five periods used in the analysis, with normalized harmonic strength for comparison, of: MDS data (green), U(0,5000) (red) and U(0,20000) (blue).
Reference


Reviewer #2:

I thank the authors for their detailed response to my review comments. While I still am somewhat skeptical of the discriminatory power of the data, I think the authors have addressed all of my queries sufficiently and the manuscript should proceed without further correction. I hope to see a follow-up article once a longer time series of mortality data becomes available.

Authors’ response: Thank you.