Reviewer's report

Title: Estimating the burden of dengue and the impact of release of wMel Wolbachia infected mosquitoes in Indonesia: a modelling study

Version: 0 Date: 18 Apr 2019

Reviewer: Giorgio Guzzetta

Reviewer's report:

In this manuscript, the authors combine several statistical and dynamical models to estimate three main quantities:

- the overall burden of symptomatic dengue infections in Indonesia;
- high-resolution (5x5km) maps of the dengue burden from different data;
- the estimated effectiveness of the release of Wolbachia-infected mosquitoes in reducing the incidence and overall burden of dengue.

The topic of the paper is of great interest for global health given the growing burden of dengue and other Aedes-borne infections worldwide, and the potential of this innovative control intervention. However, I have several questions on the methods adopted and on the interpretation and robustness of results.
Major comments:

1) National burden

* In the estimation of hospitalization rates (SI1.2), data were taken from a vaccine trial in children of age <14. If I understand correctly, the proportion of hospitalizations in this age-group was adjusted to obtain an age-independent estimate by normalizing against the relative treatment-seeking rate of the same age class in the SUSENAS data (collected for fever from any cause). Although I do not have specific suggestions on how to improve this estimate, the procedure seems a little far-stretched. The authors can show the disaggregation of age-specific treatment seeking rates in the SUSENAS data to help assess (qualitatively) the plausibility of using these data as a proxy for normalizing hospitalization rates for dengue. It is also not clear how much the burden estimates would be sensitive to alternative assumptions on this estimate.

* On a similar note, is there any support for the choice of 5-15% in the downscaling factor to compensate physician admission bias in clinical trial data?

* I could not understand how expansion factor-based methods were applied to reporting data from Indonesia; can the author explain?

* The ensemble burden estimates were computed by sampling from normal distributions parameterized to each burden estimate and Uncertainty Intervals (UI). However, judging from the position of the average, all distributions seem skewed towards the lower boundary of the UI (Fig2). Is there a statistical support for the decision to adopt a symmetric distribution for the sampling?

* Comparing national annual numbers of reported dengue cases of 144,736 (line 171 page 8) with the estimated ~1M hospitalizations (Table 1), would it be correct to say that at most 15% of all admissions to hospitals due to dengue are reported by the Ministry of Health (assuming all reported cases are hospitalized)? This seems quite low and should be discussed, also with respect to the comment above about estimating the hospitalization rates.
2) Mapping dengue burden

* Can the authors explain how do the GBM models work, what is the meaning of the model settings (complexity, learning rate, bagging fraction, shrinkage, interaction depth), and some rationale for their choices?

* Some further details are needed on the cross-validation procedure. Did the authors estimate predictive performance by leaving a subset of data (validation sets) out of the cross-validation process? Using only training/test sets may lead to information leakage and overfitting within cross-validation schemes (see for example Barla et al. 2008, Briefings in Bioinformatics).

* Fig S3 shows some assessment of the goodness-of-fit for dengue occurrence, but similar assessments would be useful also for models of incidence and seroprevalence. While I understand that a formal comparison across models is challenging, a visual information on how well each model performs on its own dataset would help evaluating how informative is each data type.

* I am afraid that the large divergence in results obtained with the three different methods make the ensemble average map too unstable for estimates that are supposed to drive the implementation of an intervention. How would the average estimated efficacy change when using individual burden maps (Fig.3A-3C) rather than the ensemble one?

* The result that 18.6% of the dengue burden may come from 15.1% of the population residing in the 10 largest cities implies that on average these cities have an incidence of $\frac{.186}{.151} \sim 1.23$ times higher than the national average, i.e. not too far from a null model assuming homogeneous incidence across the territory. Indeed, a large part of the map in Fig3D results in an almost homogeneous incidence of 2-3% with small and spatially rare deviations. Furthermore, part of the estimated excess incidence in large cities may be due to the reporting bias in cities vs. urban areas.
3) Impact of Wolbachia Programme

* What does S(a) indicate in the equations for the lambdas in model SI1.4? Should they be S_1(a), S_2(a), etc? In such case I can't figure out the reason for the decreasing exponents primary, secondary, tertiary and quaternary infection. If I understand correctly, all compartments are assumed to be subject to the same force of infection.

* Although an inflow of new susceptibles is assumed to maintain the population constant, there is no mortality term in model equations in SI1.4. so Is mortality assumed constant or age specific (and in such case, according to which data)? Given the importance of re-infections over an individual's lifetime in dengue dynamics, can the authors show a comparison of the modeled age-structure with Indonesian data, to verify that the demographic assumptions in the model are reasonable?

* It is not clear to me whether infectious compartments I1, I2, etc. include asymptomatic dengue cases; if they don't, do the authors assume that asymptomatic cases do not develop immunity? if instead they are included, do they transmit with the same infectiousness as symptomatic individuals?

* The authors declare that the 8 models used to evaluate the impact of Wolbachia programme had different structures, but equivalent parameters were extracted to be applied to the model structure described in SI1.4. I am not very convinced that this is a straightforward operation: the proportions of symptomatic and hospitalized cases (Table S3) are highly dependent on the specific transmission dynamics of a setting, due to variations in incidence and demography which in turn result into different population levels of cross-immunity and antibody-dependent enhancement. Furthermore, it is not obvious that parameters reported in other studies had the same meaning as they have in the current model. Maybe this has been already discussed and solved in reference 25, but I think some more in-depth discussion about this issue is needed.

* It is not clear to me how Beta_l and Beta_h were estimated; the authors declare that they are fit to 'unique value of hospitalized incidence', but it is not clear to me whether they used spatially varying Betas for each value of the hospitalized incidence in the dengue maps. Furthermore, I would expect the two parameters to be highly correlated so that using just one type of information would not be sufficient to identify them.

* It is very difficult to evaluate the model in SI1.5 estimating the reduction in viral transmissibility by Wolbachia infected mosquitoes: there is no information on the data used (experimental protocol, number of mosquitoes/patients tested by virus serotype, mosquito strain, DENV viremia) and neither measures of the goodness of fit nor a graphical visualization of the fit of data are shown (only the resulting estimates).
* Do the author assume identical survivals for Wolbachia-infected vs. wildtype mosquitoes? In such case, how is this supported by the experimental literature? Is the ecological mechanism underlying their ability to replace wildtype infections understood?

* The estimation of the dengue burden reduction implicitly assumes that the mosquito-to-host ratio remains the same after deployment of the intervention. However, should the abundance of the Wolbachia-infected population be higher than that of the wild-type, the number of infectious bites (i.e. the EIR) would increase overall, partially compensating the reduced transmission efficiency of mosquitoes. The authors should discuss this issue in light of results from field experiments on the replacement of wildtype mosquitoes.

* In Fig5 is shown the Wolbachia effectiveness in terms of the % reduction of cases; are these hospitalized cases only, as declared in the methods? What about the % reduction in DALY, considering that "the large numbers of self-limiting mild infections contribute more to DALY burden" (p.21, l440-441)?

Minor comments:

* p16, l338: the authors should define (in the Supplementary Material) how the Multivariate Environmental Similarity Scores (MESS) was computed, so that the reader can interpret how strict it is to require its value to be larger than zero to evaluate the "Extrapolation strength";

* How many data points were used for the splines smoothing of effectiveness curves (p11, l277-281)? Can the author show the non-smoothed curves?

* Is it correct to interpret the y axis in Figure S3 as the AUC of model predictions on test sets from the cross-validations?

* Considering that the suitability for the mosquitoes are not very important in influencing the occurrence map (Table S5 -- however it should be Table S6 since S5 is the "Detailed final reduction predictions"), did the authors attempt alternative fits with a more parsimonious number of explaining variables and check the robustness of predictions with lower-order models?

* How to interpret the covariate effects plot in S4? In addition, x-axes do not have a unit of measure.
Typos and other corrections:
* Is there a reference for l124-125 at p6?
* p9, l202: I could not find a Supplementary File 2;
* p11, l274-275: in the definition of impact the numerator and denominator are inverted;
* p13, l297-298: hard to read sentence
* Fig2 legend: there is no grey shading in the figure;
* p16, l339, FigS1 should rather be FigS3;
* p17, l365 and p21, l428 say 18.6% of incidence concentrated in 10 cities but Table 2 says 22.3%; also the figure for the land area differs (0.36% vs 0.31%);
* Fig4: legend seems to read "-10%" rather than "10%"; it's probably a residual from map cropping;
* The average effectiveness in Fig. 5 never reaches the minimum effectiveness of 26% shown in the map in Fig.6; I think it would be fair to show the full domain of predicted values; x-axis is missing a unit of measure;
* Fig6 legend: 7A and 7B should be 6A and 6B;
* Table S2: asterisk is missing in EF; furthermore, abbreviation EF (expansion factor?) is not defined in text;
* SI1.4: please check equations as per major comments above; a derivative sign is missing in all ODEs;
* In Table S3, the duration of infection for the Hopkins model is 100, I think this is a typo;
* SI1.5: please check equation of probability of onward transmission.
Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

No

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

No

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

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 have been a consultant for Merck on a one-day meeting on varicella-zoster. I am currently involved in a research project funded by Merck on varicella-zoster. I have no other competing interests.

**Statement on potential review bias**

Please complete a statement on potential review bias, considering the following questions:

1. Did you co-author any publication with an author of this manuscript in the last 5 years?

2. Are you currently or recently affiliated at the same institution as an author of this manuscript?

If you can answer no to all of the above, write 'I declare that I did not publish with these authors in the last 5 years and also meet the affiliation criteria'. If your reply is yes to any, please give details below.

I declare that I did not publish with these authors in the last 5 years and also meet the affiliation criteria

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