Author’s response to reviews

Title: Estimation of cumulative number of post-treatment Lyme disease cases in the US, 2016 and 2020

Authors:

Allison DeLong (Allison_DeLong@brown.edu)
Mayla Hsu (mayla.hsu@globallymealliance.org)
Harriet Kotsoris (hokpapa@gmail.com)

Version: 2 Date: 11 Dec 2018

Author’s response to reviews:

Dr. Anna Bajer, PhD habilitation
BMC Public Health

December 11, 2018

Dear Dr. Bajer,

Thank you for allowing us to once again revise manuscript reference number PUBH-D-18-0271, “Estimation of cumulative number of post-treatment Lyme disease cases in the US, 2016 and 2020.” We appreciate the thoughtful reviewer comments, and have used them to improve the manuscript. Edits on the revised paper are highlighted with track changes. We have provided a point-by-point response to Reviewer 1, who had remaining comments, and indicated where these can be viewed by stating the page and line number on the edited manuscript.

We hope that the paper is now suitable for publication in BMC Public Health.

Sincerely,

Allison K. DeLong
Brown University
Technical Comments:

(1) Please include an email address for each author on the title page.

Response: These are included and now fit on the first page.

Allison DeLong’s email: Allison_DeLong@brown.edu

Mayla Hsu’s email: mayla.hsu@gla.org

Harriet Kotsoris’ email: hokmd@aol.com

(2) Please include an 'Acknowledgements' section in your Declarations.

Response: In line with publications from this journal, we have moved the Acknowledgements to the first statement in our Declarations section.

(3) Please remove the figure title/legend within figure file 1.

Response: We have removed the “A” and “B” legend from within Figure 1

Reviewer reports:

Gerrit Gort (Reviewer 1): The authors improved the paper substantially, taking the comments into careful consideration.
I like the addition in the Discussion section about two types of uncertainty: uncertainty in disease incidence and treatment failure rates handled by different "settings" (where all six settings are more or less realistic now) and uncertainty, given a specific setting, stemming from probability distributions, which would give rise to the data.

Reviewer Comment: In the Conclusions it should be mentioned that the variability in the estimates is very high, as mentioned in Results. E.g. a range of 65,000 - 1,500,000 is very large indeed.

Response: We have added “were highly variable and” to the 2nd sentence in the Results section (page 15, line 15-21).

The sentence now reads, “The 2016 estimates were highly variable and ranged from a low of 69,153 persons (95% CI 51,891 to 89,196) with Scenario A and failure rate of 10%, to 1,523,754 (95% CI 1,280,832 to 1,811,412) with Scenario C and failure rate of 20%.”

Reviewer Comment: A few points remain feasible, in my opinion.

- Contrary to the addition in the Discussion part (see above), I think that disease incidence and treatment failure rates are currently not treated in a similar fashion. Both are indeed handled in different "settings", but the treatment failure rates get an extra layer of uncertainty through the beta distribution: simulation of the occurrence of PTLD does not come from a binomial distribution with 10% and 20%, but from a binomial distribution with probability of success sampled from a beta distribution. The same procedure could be used for disease incidence: starting from a specific setting for the expected number of PTLD cases in a year a random deviation (expressing the uncertainty about it) from that number could lead to a count, to be used as parameter for the probability distribution (like Poisson) for counts. However, the parameter for the count probability distribution is the expected number of PTLD cases directly.

- Again, there is little rationale for the use of the Poisson distribution for counts, which lead to very small relative variation. Other distributions for counts exist, which allow for more variation, like the negative binomial distribution (but admittedly without much rationale too). The differences between the settings, however, will probably drown whatever variability would be obtained from the probability distributions.
Response: We agree that we have been treating the variability for the failure rate and disease incidence differently and have modified our methods to incorporate the reviewers comment to be more consistent. As the reviewer suggested, we have used the negative binomial distribution to simulate the count of new infections each year, instead of the Poisson distribution. The mean of the negative binomial distribution is set to the expected value based on the incidence scenario, and the variance is set to the mean multiplied by an overdispersion parameter. We estimated the overdispersion parameter by fitting overdispersed Poisson regression models to the CDC surveillance data for 1997 to 2017. CDC reporting criteria changed in 2008 and we keep these two timeseries separate. Assuming linear growth over time for years 1997-2005 (as assumed in Scenario A) overdispersion is estimated to be 224; for years 2008-2017 overdispersion is estimated to be 349. To be consistent with both scenarios, we set overdispersion to 320. While we expected the mean of our simulations to remain the same, we expect the confidence interval to widen. We have updated Table 1 and the manuscript text with the new results.

We change the text in the Methods section to reflect this modification. Specifically, we replace reference to the Poisson with the negative binomial distribution in the first paragraph in the Methods section (page 6, lines 31-39). The sentence now reads: “We base our simulations on the technique presented by Crouch et al. (11) and utilize the probability distribution function from classical statistics that most closely represents the type of data used in each step of the simulation, i.e. the binomial distribution for “yes/no” data and the negative binomial distribution for overdispersed count data.”

The following sentences were changed in the Methods section (page 9, lines 32-54 to page 10, lines 11-13):

“Variability: While the mean (or expected number) of new infections was presented as input, the simulations allowed the actual number to vary stochastically using the negative binomial distribution with the variance set to the expected value multiplied by an overdispersion parameter. This allowed uncertainty to increase with a higher expected number of new infections. The dispersion parameter was estimated by fitting overdispersed generalized linear Poisson regression models to the number of confirmed cases in the CDC surveillance data from 1997 to 2017, using linear growth over time as the only independent variable (20). We fit two regressions due to the assumption of Scenario A of linear growth from 1997-2005 with constant
incidence thereafter and because the CDC employed two different reporting criteria before and after 2008. The dispersion parameter estimated from the 1997-2005 data was 224, and the dispersion parameter estimated from the 2008-2017 data was 349. To be consistent with both models and err on the side of higher variability, the dispersion parameter was set to 320 in our simulations.”

Also, the mean for Scenario A changed slightly as we noticed an error in the R code for incidence scenario A. We planned to allow for linear growth from 1980 to 2005 with constant mean incidence thereafter, but the code accidentally calculated linear growth from 1980 to 2008. This has been corrected in the new output in Table 1 and the manuscript text.

Reviewer Comment: p7, line 12-15 Please give reference for this (PTLD death rate equals general US population death rate), or mention that you explain this later, as you do give a reference at p 9, l 24 in the survival part.

Response: We have added to page 8, line 31-32, “Since LD is rarely listed as a cause of death (ref 22)” to the sentence, “we assume death rates for those with PTLD are identical to those of the general US population.”

For increased clarity, we changed one line in the Discussion, page 17, lines 51-53. Rather than vaguely writing “mixed conclusions”, we wanted to specifically indicate lack of sustained improvement after prolonged antibiotic therapy. The sentence now reads:

“”However, studies evaluating long-term, or repeated, antibiotic treatment of PTLD patients have not shown sustained improvement (35-38), although there is evidence some subgroups may benefit from retreatment”

Claudia Coipan, Ph.D. (Reviewer 2):

Reviewer Comment: I was pleased to read the revised version of the manuscript entitled "Estimation of cumulative number of post-treatment Lyme disease cases in the US, 2016 and 2020". As I mentioned before, the results of this study provide a useful metric in calculating the
disease burden associated with sequelae of Lyme infection and will be of interest for the readers of BMC Public Health.

The revised manuscript addresses the questions raised in the review process; the methodology has been adjusted and the results presented accordingly. I have no further comments regarding the manuscript.

Response: Thank you.