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: 1 Date: 05 Nov 2018

Author's response to reviews:

Dr. Anna Bajer, PhD habilitation
BMC Public Health

November 4, 2018

Dear Dr. Bajer,

Thank you for allowing us to revise manuscript reference number PUBH-D-18-0271, “Mathematical determination 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 each reviewer and editorial point raised and describe exactly which amendments have been made to the manuscript text and 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
Technical Comments:

1. Please include an email address for each author on the title page.
   
   Response: We have included email addresses for each author on page 2, line 12.

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

   Response: We have included an Acknowledgements section in the Declarations. We ask your permission to add this “Disclaimer” to the Declarations section: “The views and opinions expressed here are those of the authors and do not necessarily reflect the official policy or position of Brown University or Global Lyme Alliance.”

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

   Response: We have removed the figure title from within Figure 1.

Reviewer #1 Comments (numbering added):

1. The outcomes of your simulations depend heavily on the 329,000 new cases per year, which are based on a health insurance claims database (P4). Are you sure that these 329,000 are really new cases, and not duplicate diagnostic tests (e.g. for second opinions) or replicate tests for earlier diagnosed cases? If that would be the case, the number of actual new cases may be substantially lower. In other words, how reliable is this rate?
Response: The authors of the article from which we drew this figure did in fact try to control for this, defining “incident diagnosis” as one in- or out-patient event per year. A patient with more than one of these events within a year was counted simply as one diagnosis. To clarify this point, we have added the word “incident” to the manuscript, page 6, line 27, in the sentence “Nelson et al. (2015) used data from a health insurance claims database to estimate there have been approximately 329,000 incident Lyme diagnoses per year during 2005-2010 (range 296,000-376,000).”

2. To continue with the 329,000: in your simulation you assume Poisson variation. Poisson variation is, relatively speaking, very small for such large numbers: a Poisson distributed variable with mean 329,000 has variance 329,000, so standard deviation \( \sqrt{329,000} \approx 566 \). Therefore, you will sample counts with a 2SD range of \( 328,000 \pm 2 \times 566 \approx 327,000 - 329,000 \). That is not expressing a lot of uncertainty about this number. Would a larger variance not be much more realistic, if it is to express uncertainty about the number of yearly new cases of LD?

Response: This is a very valuable and important comment. Our response to it attempts to address several of this reviewer’s other comments as well. This question of overall uncertainty has nagged at us from the start of the project and we eventually chose to address this explicitly, by using 6 settings (3 incidence scenarios and 2 treatment failure rates). Although these incidence scenarios are based on values presented in the literature, certainly one of these scenarios is closest to the unknown true epidemic. We are confident that future research will help uncover a more precise quantification of the Lyme disease epidemic that can be used in similar calculations.

To better clarify this limitation and our approach, several sentences were added to the first paragraph in the Methods section and a new section was added to the DISCUSSION.

In the Methods, page 5, lines 35-45, we add:

“The six settings represent three scenarios for LD incidence and two treatment failure rates. While any one of the settings may currently be considered more realistic, more research is being conducted that may change our understanding. At that time, an improved framework could be developed that incorporates all uncertainty into one simulation.”

And the new section in the Discussion, page 16 line 37 to page 17 line 9, is:

“We acknowledge and incorporate uncertainty in disease incidence and treatment failure rates by simulating six settings, each providing an estimate of prevalence under the assumption that the setting is correct. As a result, estimates of uncertainty come from the probability distribution function generating the type of data used as input, and do not incorporate uncertainty about the
mechanism generating the data (i.e. the setting). The uncertainty in the data-generating mechanism is shown by examining the differences in the results across all six settings.

An alternate approach could have been to incorporate all uncertainty within one scenario, perhaps weighing each scenario based on some knowledge base. The result would be essentially a weighted average of the results presented here with very wide 95% coverage intervals. Alternatively, this question could be cast in a Bayesian framework and could incorporate expert opinion as priors. Unfortunately, as of now, the weights and priors are unknown. An improved diagnostic test, with national surveillance, and research into treatment failure rates will likely provide more precise information to indicate which setting optimally fits the dynamics of the epidemic.”

3. The authors work with two separate treatment failure rates (10% and 20%; P5 top), allowing deviations from these numbers in the (separate) simulations, based on beta distributions. But the numbers 10% and 20% indicate that is a lot uncertainty about the failure rate. Why do you take this approach, and not work with a single, but highly uncertain failure rate?

Response: We place the uncertainty about treatment failure in a similar category as the uncertainty about disease incidence and our decision to evaluate 6 settings. Therefore, the response to this comment is provided above in our response and edits to your Comment 2. We have also provided further description in the Methods section (under Treatment Failure), page 6 line 55 to page 7 line 7, as to why the true failure rate is unknown:

“Given the variability of treatment failure due to regional, geographical differences, socioeconomic factors, co-morbidities, treatment delays, and non-standardized treatment protocols, we chose to encompass both extremes of this range, basing our estimations on either 10 or 20%”

4. The authors try three different scenarios (Pp 5-6). Scenario A is VERY unrealistic: a gradual increase from 0 to 30,000, and then suddenly an enormous jump to 329,000 new patients per year. Why do you include such an unrealistic scenario?

Response: In the first submission, we included scenario A as our most conservative scenario because it allowed for the lower number of surveillance cases reported to the CDC until the more recent publications have indicated actual incidence is likely 10 times higher. We agree our original scenario A as a whole is unrealistic and propose an alternate variation of Scenario A in this resubmission. A recent article based on Lyme disease surveillance found that an average of 34,449 cases were reported in the years of 2005-2010. Therefore, in our revision, we suggest that Scenario A be composed entirely of surveillance cases and be modeled by linear growth of the epidemic from 0 cases in 1980 (beginning of the epidemic) until 2005, when the available surveillance case numbers were reported as 34,449 cases. After 2005, the revised Scenario A
would have annual case numbers remain stable at 34,449. This revision has been added into the appropriate section of the methods section with resulting changes in the results and discussion sections and we have added the pertinent reference (Schwartz et al). This change also addresses a comment from Reviewer 2.

The revised sentence in the Methods, page 7 lines 37-45, is: “Scenario A represents LD cases captured for surveillance purposes and assumed linear growth from 0 cases in 1980 to 34,449 cases in 2005 and remained stable at 34,449 annual cases from 2005 onward. These values derive from surveillance data published by the US CDC (19).”

5: The authors assume that survival rates for patients with PTLD are the same as for the general populations (P6 bottom). How realistic is this? Can you supply a reference for this assumption? May be PTLD patients experience high mortality due to other causes which become lethal in their situation.

Response: The survival rates for patients with PTLD were assumed to be the same as for the general population. We have provided more justification for this assumption and added two references in support of it. The following sentence was modified in the Methods section, page 9 lines 19-27, and the relevant references added below:

“Mean: While nine cases of fatal Lyme carditis have been recognized in national surveillance data from 1985-2018 in the US (21), a review of death reports and death certificates in the US from 1999-2003 cited Lyme disease as a rare cause of death (22). Survival rates for patients with PTLD were, therefore, assumed to be the same as the general US population and survival rates after 2014 were assumed to be as in 2014.”

Prevention CDC. https://www.cdc.gov/lyme/signs_symptoms/lymecarditis.html 2018


6: The deterministic estimate of prevalence (P8) is not "validating our results" (P4) in a general sense, but only as a check for the simulations. I think the wording on page 4 is too strong.

Response: Accordingly, we have changed this sentence in the first paragraph of the Methods and added the following sentence to this paragraph, page 5, lines 33-36: “We also use a simple deterministic approach as a check for the simulations.”
7: In the discussion (P11, bottom) you mention that the expected number of new LD cases is critical.

Response: We agree, hence the scenarios. Please see response to comments 2 and 3 above.

8: An alternative approach to use could lie in Bayesian statistics (although I cannot see the exact workings for it directly here), as it incorporates all sorts of uncertainty to arrive at a posterior distribution for the quantity of interest. May be you can make some reference to this in the discussion.

Response: This is a very good suggestion and we have referred to this in the Discussion section in our new paragraph with the sentence on page 16, lines 57-61:

“Alternatively, this question could be cast in a Bayesian framework and could incorporate expert opinion as priors.”

Reviewer 1 Smaller remarks:

Comment: P2 L22: (and others) You use the word "calculated" in many places. For me a better word is "estimated": you obtain an estimate for quantities (namely, the Nth year prevalence of PTLD) with related uncertainty.

Response: We agree with this observation and have replaced our use of “calculated” with “estimated” at this line in the abstract, lines 23 and 42, and throughout the manuscript. We have also changed the title of the manuscript to “Estimation of cumulative number of post-treatment Lyme disease cases in the US, 2016 and 2020.”

Comment: P2 L22-24: the cumulative numbers of □ the prevalence of?

Response: We have replaced “cumulative numbers” with “prevalence of” in the abstract, line 26.

Comment: P3 L43-47: You mention here that the "precise societal burden …has never been adequately quantified". And a bit further : "To address this absence…". This suggests that your estimation is giving a quantification of the precise societal burden. Do you really think that this is the case? Or are you just adding a small contribution?

Response: We agree that this series of statements likely overstates the contribution of our paper in that we do not actually quantify the precise societal burden of Lyme; however, providing a
statistical framework to estimate prevalence of PTLD and providing some estimates of the prevalence of PTLD in the US is an important contribution in the process of determining the societal burden of PTLD. To our knowledge, such an estimate has never been attempted nor published.

To clarify our contribution, we modifying the pertinent sentences in the Background section of the manuscript, page 4 line 53 to page 5 line 7, which now reads:

“As a critical first step in addressing this absence, we developed a statistical framework to estimate the prevalence of PTLD in the US by using Monte-Carlo simulation techniques and applying it under six settings representing various assumptions about the course of the US Lyme disease epidemic. The settings make use of available published data on the US Lyme disease epidemic and demonstrate the wide range of measures provided in the literature.”

Comment: P3 L51: you mention here "mathematical modeling". I see more statistical modeling, as you use probability distributions (like binomial, multinomial, uniform and beta) to quantify uncertainty.

Response: In an effort to better clarify our methodology, we have replaced “mathematical modeling” with “simulation techniques” or “statistical simulation techniques” in the Background and Conclusion sections, specifically page 4 lines 55-58, page 16 line 6-7, and page 17 line 50-51.

We have also changed the title of the manuscript to “Estimation of cumulative number of post-treatment Lyme disease cases in the US, 2016 and 2020.”

Comment: P4 L6: in what sense are the inputs "conservative"? On the low side or on the high side?

Response: Our goal was to base our inputs on estimates provided in the peer-reviewed literature. For example, while the 10-20% failure rate is commonly reported and most clinicians and researchers admit it is at least this high, there are published estimates up to 40% among specific patient subgroups (e.g. those with the longest treatment delays). Not using these extreme values
was our attempt to be conservative. To be more precise as to the methodology we employed, we have removed the ambiguous use of “conservative” and edited this sentence, page 5 line 20-23, to now read:

“To estimate prevalence, we used inputs on Lyme disease incidence and treatment failure rates commonly reported in the peer-reviewed literature.”

Comment: P4 L9: please specify why you think that your statistical assumptions are "appropriate"; sometimes I have some doubts (e.g. with Poisson distribution, see earlier).

Response: The use of “appropriate” here was to acknowledge that the Poisson distribution is the correct probability model to model count data and that setting the variance equal to the mean is appropriate for the Poisson distribution. To clarify this, we have added additional details to the first paragraph in the Methods section, page 5 line 19-34:

“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 Poisson distribution for count data.”

Comment: P5 L22: it may be good mention here already that you assume that death rates for PTLD patients are identical to these from the general population.

Response: We have added this to the pertinent section on page 7, lines 12-15:

“We assume death rates for those with PTLD are identical to those of the general US population.”

Comment: P5 L 33: mention that this is about incidence of LD (and not PTLD)

Response: We have replaced “Incidence” with “Incidence of Lyme disease infections” at this line on page 7, line 27-28.

Comment: P5 L35: the word "arguably" is essential here, because scenario A is completely unrealistic (see above).
Response: Hopefully the modifications we have made to scenario A and clarifications we added in the Background, Methods and Discussion will address this comment. Please see our response to Comment 2 above.

Comment: P5 L53: would exponential growth ever be a realistic option in the expanding phase of a disease like LD, which is non-contagious? If not, it doesn't make much sense to make the comparison here, suggesting that linear growth is highly conservative.

Response: Indeed the growth rate of the epidemic is uncertain, due to the paucity of data available on disease incidence at various time points. We have added the following edits to this section, page 7, line 51-56, to reflect this uncertainty. “Use of linear growth in our predictions is conservative over exponential growth, the latter a potentially realistic option in an expanding epidemic, due to rapid growth of the vector population and tick habitats.”

Comment: P6 L40-49: The part about the use of the beta distribution is a bit unclear. May be you can give an example, to give the reader in impression of the uncertainty you introduce here (e.g. with p=0.10, you sample a p from beta(50,450) which has mean 0.10 and SD 0.013, so roughly 2 SD range for p is 0.074-0.128).

Response: We agree adding in additional details is beneficial. Therefore we have added the following sentence to this section on page 9, lines 6-12:

“For a failure rate of 10%, p=0.10, the mean is 10%, and the 2 standard deviation (SD) range is 7.4 to 12.8%; for a failure rate of 20%, p=0.20, the mean is 20%, and the 2 SD range is 16.4 to 23.6%.”

Comment: P9 L11: explain how you calculate a "CI”. I guess CI stand for confidence interval, but note that confidence intervals in statistics have a special meaning. Here you produce the 2.5 and 97.5 percentiles from the 500 simulated results, aren't you?

Response: The reviewer is correct, our 95% Coverage Intervals (CI) are taken as the 2.5th and 97.5th percentiles of the, now, 5,000 simulations. This sentence, page 10 lines 17-24, now reads:

“For each simulation, we performed 5,000 runs and provide and provide the median and the 2.5th and 97.5th percentiles as coverage intervals (CI) of the results.”

We also made the appropriate modification to the heading of Table 1 (page 21, lines 4-12):
“For the simulation, results are presented as the median and the 2.5th and 97.5th percentiles as coverage intervals (CI) over 5,000 runs.”

Comment: P9 L40-42: Figure 1 is not showing that "the relative distribution by age and gender was insensitive to failure rate", because you show the lumped results from the simulations for 10% and 20% (I guess).

Response: Thank you for pointing this out. We have clarified this sentence on page 12, lines 44-47, to now read:

“However, while the prevalence of PTLD varies by age and gender (Figure 1), the relative distribution by age and gender was similar across all scenarios (not shown).”

Reviewer 2 Comments:

1. Throughout the manuscript, the subjective and variable nature of late Lyme disease is underlined by the authors, which I think is completely justified. Also, the authors name the unreliability of the diagnostic tests in late phases of the disease. Yet, the tendency is to deem the tests as lacking sensitivity and therefore to trust more the clinical diagnosis - as described in Nelson et al. (2015). My point is, couldn't the laboratory tests also be lacking specificity, and couldn't the clinicians diagnostic be subjective. Given the fact that ticks carry more potentially pathogenic microorganisms, and not only Borrelia burgdorferi, couldn't it be that the high numbers of PTLD are actually an overestimation?

Thus, I consider it objective to include in the analysis also a scenario using the official estimates based on the direct surveillance (30,000).

Response: Dr. Coipan raised an excellent point that basing our estimate on test results or clinician diagnoses may be limiting and suggested including a scenario based on surveillance data. Reviewer 1 also had concerns about scenario A. Accordingly, we have changed Scenario A in our analysis to be based on the number of surveillance cases with linear increases in Lyme disease incidence from 0 cases in 1980 to the surveillance-derived number of 34,449 cases in 2005 and constant 34,449 thereafter.

Please see our response to Reviewer 1, Comment 4.
2. How was the number of runs for the simulations chosen? What was the uncertainty you were willing to accept so that the chosen number of runs was 500 and not higher? Should one be expected to trust the results of the simulations, these details are essential.

Response: We originally chose 500 runs per simulation due to the long time it took for each simulation (200 runs/hour). We have now increased the number of runs per simulation to 5000 which should provide more accurate estimates of the median and 2.5th and 97.5th percentiles.

3. For reproducibility of the analyses you should indicate what program/software you used for performing the analyses

Response: We have added the following reference and sentence to the Methods section, page 10, lines 21-24, of the manuscript: “All analyses were performed in R version 3.5.1.”


4. Inaccuracy (paragraph 2, page 11) is a very broad term and also inaccurate when it comes to tests - there you have sensitivity and specificity. Could you make it more clear for the reader what do you mean by that?

Response: The Lyme disease blood test that is accepted by the mainstream medical community and CDC is the two-tiered test, and its inaccuracy is due to low sensitivity during acute infection, and its narrow specificity for limited bacterial strains that cause Lyme disease. We have included this in the referenced section, page 15, lines 33-37, in the following edit. “Its inaccuracy, due to low sensitivity in early infection and inability to capture bacterial strain variation, is overall a dismal 50-60%”.