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: 3 Date: 14 Jan 2019

Author’s response to reviews:

Dr. Anna Bajer, PhD habilitation

BMC Public Health

January 14, 2018

Dear Dr. Bajer,

Thank you for allowing us to 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 are very appreciative of Dr. Gort’s additional comment and advice. I have be very impressed with his reviews and know this paper is much improved because of him. We have provided a response to the one remaining comment and have indicated where changes can be viewed.

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

Sincerely,

Allison K. DeLong
Reviewer reports:

Gerrit Gort (Reviewer 1): The authors made a good effort to incorporate the comments. Currently I have no remaining comments, but one technicality related to the negative binominal distribution. The authors sample counts of yearly new infections using the negative binominal distribution, which is a more reasonable distribution than the Poisson distribution here. For the negative binomial distribution an "overdispersion" parameter is employed. This overdispersion parameter is estimated using the CDC surveillance data, using an overdispersed Poisson model. I suspect that this approach is not correct, because a standard overdispersed Poisson model indeed has a constant overdispersion parameter (with variance equal to the mean multiplied by overdispersion parameter), but not so for the negative binomial distribution: its variance is $\mu + \mu^2 / size = \mu*(1+ \mu/size)$ (using the parameter names from R in function nbinom). From this you see, that the multiplication factor changes with $\mu$. The parameter $size$ is the parameter related to overdispersion (relative to Poisson). My suggestion is to use the CDV surveillance data, using a negative binomial model (you can use the R function glm.nb (from mass package), to estimate this parameter. Next, use this parameter to sample from negative binomial distribution(s).
Author Reply:

We are very appreciative for this constructive comment and have done as Dr. Gort suggested. We modified the shape parameter used to simulate incidence using the negative binomial. Our negative binomial fits (using glm.nb from the “MASS” package in R) to the CDC data, resulted in a dispersion parameter of 112 using CDC confirmed cases 1997-2005 and 127 for total cases from 2006-2017. We used the average of these, 120 in our simulations.

The methods have been modified. The paragraph (page 6, lines 27-55) now reads:

“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 mu + mu2/size, where mu is the mean and size is 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 negative binomial 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 (19). 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 112, and the dispersion parameter estimated from the 2008-2017 data was 127. To be consistent with both models, the dispersion parameter, size, was set to 120 in our simulations.”

In addition, Table 1 is updated and the numbers used in the abstract, results and discussion sections have been modified with the new values in Table 1.