Article details: 2021-0024	
Title	Estimating population health burden of lyme disease in Ontario, Canada: a microsimulation modeling approach
Authors	Stephen Mac MBiotech, Gerald A. Evans MD, Samir N. Patel PhD, Eleanor M. Pullenayegum PhD, Beate Sander PhD
Reviewer 1	Dr. S.K. Peasah
Institution	
General comments (author response in bold)	Page 4.1. can you please rewrite the objective as "to estimate the health burden of Lyme disease using microsimulations…?" Instead of "to develop an individual-level state-transition model to estimate". It more accurately represents your objective. **We have revised the objective as suggested in the Introduction (p.3).** Could you please expand on your rationale for using the following values to clarify your methods? Mean age of 37.62 years. Is it the mean age of Ontario residents or you had another rationale for it? **Similar to the above comment, this is the mean age of people in Ontario, and is used to create the distribution from which each individual in the model's age is sampled from. We have added this in the Methods (p.4).** High-risk areas. Any reason to exclude low-risk areas? Insignificant? **Similar to the above comment, we did not exclude low-risk areas and have attempted to clarify that 63% of the population lives in high-risk areas in the Methods (p.4).** Since death is rare in Lyme disease- is the life expectancy given (81.02), that of Ontario residents or people with Lyme disease? Same for why it was not included in the Markov model. **The life expectancy is that of the Ontario population. There are no studies that measure age at death for Lyme disease patients, and because this was a population-based study, we used the life expectancy of Ontario population. Revised in Methods (p.4).** Could you add the rationale for using 1.5% discount rate? **See comment above to Editor re: discounting.** Could you give a reference for the ranges given to distinguish between "early localized, early disseminated, and late disseminated? **We have provided a reference for these time periods from the Government of Canada in the Methods (p.5).** Could you expand on the role of the probability of hospitalization and length of hospital stay in the model? **The role of hospitalization and length of stay has an effect on the quality-of-life (utility) of the person depending on whether or not they develop sequelae. The hospitalization is relatively short and so the effect is minimal. We added a brief statement in Methods (p.6).** Interpretation. Change "number of individuals" to "percentage of individuals" Page 9.19 **This has been revised in the Interpretation (p.9).**

	Although this is well written there are some grammatical issues worth correcting throughout the manuscript. **Thank you for this comment, we have revised throughout where we found grammatical errors.**
Reviewer 2	Dr. Affan Shoukat
Institution	Yale University
General comments (author response in bold)	Despite being an individual based model, the authors have overlooked to provide uncertainty bounds for their results of point estimates. Instead of point estimates, the authors may want to identify a range of possible values the outcomes of interest might take. Additionally, it is unclear how the authors incorporated the ranges for parameters described (Table 2) in the simulations. To establish robustness and provide uncertainty bounds on their results, the authors may want to consider random sampling techniques. For example, a number of stochastic simulations can be performed. where in each simulation parameter values are sampled from their relevant distributions. This will allow authors to provide a point estimate (i.e. mean or median) as well as a range around this estimate (i.e. CI or IQR). **Thank you for this comment. Similar to the above comments, we have run 100 simulations of 100,000 individuals. We have included the uncertainty along with the results throughout the manuscript.** **Regarding the ranges, we show these in Table 1 as plausible or full range of the parameters that are used in the model. The base-case analysis uses the base-case value, but the range is used for deterministic sensitivity analysis. We have clarified this in the Methods (p.7).** The sentence "The likelihood of clinical diagnosis after presenting with EM in high, and low exposure areas was 58%, and 26%" in the results is unclear. Are the authors discussing the sensitivity of the diagnostic test that confirms the diagnosis? **We are discussing the likelihood someone is clinically diagnosed (without a test) in high-risk and low-risk areas for Lyme disease by their physicians using case definitions. Clinical diagnosis is different from laboratory diagnosis. We feel this is already clear in the statement and have decided not to revise.**
Reviewer 3	Dr. John Webb
Institution	Cumming School of Medicine, University of Calgary
General comments (author response in bold)	I admire the attention to detail that has gone into modelling the natural history of the disease and how that may vary, depending on the interaction with the health care system. **Thank you for this comment.** I would like to see a more clear and compelling statement of the reason for the study. I understand the rationale for your study to be that it offers a more nuanced and accurate modelling of the burden of disease than other studies; but those studies also sought to answer questions of cost-effectiveness of vaccines or oral medication. Is there a question your study will answer? If it is primarily intended to offer better modelling of disease than found in those studies, why is that important? Since Lyme is already a nationally reportable disease, and an acknowledged public health problem with a high media presence, why do we need to further characterize the burden of disease? I would suggest there is no need to

	raise awareness, but I believe there are new vaccines in development. I wonder if providing this context might clarify the value of modelling the burden of disease. **Thank you for this comment. Similar to the comment above where we addressed why we are using a modelling approach to assess health burden of Lyme disease, the question we sought to answer was the population-based health burden associated with Lyme disease in terms of QALYs. Previous models were not built to answer population-based health burden as they were interested in answering other questions and therefore did not include many aspects (diagnostic, various treatment, hospitalization, post-treatment Lyme disease syndrome) that are important in the detection and management of Lyme disease. We elaborated briefly but it has been mentioned in the Introduction (p.3).** **This model can also be used in the future to answer questions of cost-effectiveness of vaccines or public health interventions since it is a population-level model that incorporates the entire disease history from infection (tick bite) to death. We have alluded to this and have strengthened this part of the Interpretation (p.12-13).** **While Lyme disease is already a national notifiable disease meaning it is reported and has media presence, the actual burden of the disease is not well understood. One of the most common complaints amongst those infected with Lyme disease is that their quality-of-life is not considered as health burden. An epidemiological study on long-term outcomes in Canada does not currently exist (although it is underway with the Canadian Lyme Disease Research Network, that was funded through a Federal Framework for Lyme Disease Act).** **Similar to Long-Covid that is of attention now, there have been those with PTLDS or "Chronic Lyme disease" that report long-term symptoms post-treatment. Our model does not suggest that there is a need to raise more awareness (although that is a possible intervention), but rather there are considerable QALYs lost resulting from Lyme disease. We feel that capturing this population-level burden can help shape policy and decision-making in the future when evaluating interventions. We have tried to emphasize this context more in the Interpretation (p.13).** I noticed there is no section on model validation, although you do say in the interpretation section that "the number of individuals developing PTLDS was similar to the 10% reported in the literature." Could you add subsections on external validation in the methods and results sections, describing other outputs of your model (perhaps drawn from the base case) and how they compared to published or administrative data? I see you did attempt to compare the QALYs to HALYs from other studies, but perhaps you could compare some of the other figures in the base case table? **Please see comments above re: validation.** In your Markov diagram, I suggest changing the title of the "Infected" bubble to something along the lines of "Undiagnosed infection" to make it easier to understand. At first look, I was confused how an individual could progress from

	being Healthy to having Early Local Disease without going through the Infected stage. After I read the text, I understood that Infected actually meant the person was infected but undiagnosed. I would find the diagram easier to understand if the bubble said that somehow. **Great point, we have changed this in Figure 1.** I don't believe your model allowed for the possibility of antibiotic-refractory arthritis. This might not change your results, but it might be appropriate to explicitly provide a rationale for not modelling this infrequent complication **Similar to the above comments, the data that is available to us does not provide probabilities of specific sequelae. Overall, it should not impact the results so long as the health state utility values are capturing the type of sequelae.**