

Dr. Benjamin Bolker
Professor
Departments of Mathematics & Statistics and Biology

1280 King Street West
Hamilton, ON L8S 4K1 Canada
phone (905) 525-9140 x23320
fax (905) 522-0935
e-mail bolker@mcmaster.ca
February 11, 2017

To the editor,

We are happy to have the chance to revise our manuscript (and thank you for your patience while we've taken a little longer than average to do so). Besides general clarification, some of the main changes are that we

- added a figure illustrating the main model structures we use
- switched our primary descriptor of viral evolution back to SPVL, as recommended by reviewer #1
- implemented vital dynamics and heterogeneity in conjunction with each of our four 'base' models (i.e., models with and without instantaneous switching, with and without extra-pair contact); the detailed results for these models are given in the supplementary materials, but we do comment on the results in the main text.

We look forward to your decision on the revised version.

Reviewer #1

The outstanding issue is that the paper remains quite challenging to read. It is just not written as clearly as it could be. In many paragraphs, the authors launch into discussion of complex points with long, multi-clause sentences, instead of explaining their motivation and results in simple, concise phrases. Too often the authors do not make it clear WHY they did something. Reading the original papers by Fraser's group on this topic (PNAS 2007 and PLoS Comp Biol 2011) provide good examples of clear writing.

This is a very general comment, but we have gone through the entire manuscript and tried to clarify it.

Abstract: *The first few sentences of the are confusing, especially "Changes in the fitness landscape generally select for higher virulence early in an epidemic". Do you really mean to suggest the fitness landscape - which is the mapping between parasite genotype and phenotype - is dynamically changing? And do you think that the main driver for changes in virulence is simply the stage of the epidemic? If so, then it would be better to complete this sentence with a brief description of that idea and the expected trends in virulence over time. The other important factor here is that evolution of virulence is mostly likely to occur after a pathogen is introduced into a new host. It may have been optimally adapted to the original host, but differences between hosts mean it may no longer be optimal. This seems very relevant in the HIV case, as it is a zoonosis from SIV.*

We have modified the first few sentences to make them more concrete and (hopefully) more understandable. Yes, we did mean that the stage of the epidemic is a strong driver of virulence evolution, and the one we focus on here. We have removed the potentially confusing term "fitness landscape".

The point about initial virulence being measured immediately after a species jump is interesting, but seems peripheral to our goals for this paper.

Author summary: *In the first sentence, I think it would be clearer to say "infectivity" instead of "prevalence". Obviously disease prevalence changes over the course of an epidemic .. the epidemic starts with one individual being infected, and then rapidly increases. This has nothing to do with evolution - it would occur for any strain with $R_0 > 1$*

as it spreads to others, even if no evolution could occur!

We were previously trying to pack too many ideas into a few sentences. We have made the text more concrete, clarifying that we do indeed mean to describe the fact that prevalence changes over the course of the epidemic, and that this changes the evolutionary environment for the pathogen.

Introduction: *While the intro has been improved since the last version, it is still relatively confusing. It would really be improved by explicitly and concisely explaining the goals of this study, and by addressing the other concerns below.*

Second sentence - would be much clearer to say something like "... pathogens with higher reproduction numbers - that is, the number of secondary infections caused by a single infected host over the course of its infectious period - will tend to increase in prevalence relative to strains with lower reproductive ratios." Your definition of "reproduction" was not very clear, and it's not a good idea to call R_0 a rate, as it does not have units of time.

We were indeed thinking of reproduction rates (i.e., r) rather than R_0 here; we may be slightly more attuned to r because of our eco-evolutionary work, but we agree that it's clearer to start out with an R_0 framing, which will be more familiar to technical readers (and clarifying the r - R_0 distinction is really too much to try to pack into the first few sentences of the introduction).

Confusing sentence "... in studies of discordant couples (...), HIV virulence as measured by the rate of progression to AIDS was both heritable and covaried with the set-point viral load (...) and the probability of transmission". - You are combining results of multiple different studies here. I don't think there is a study that compared rate of progression to AIDS with probability of transmission. I think studies have a) compared SPVL to rate of progression, b) showed that SPVL was heritable, c) compared SPVL to probability of transmission. I think b) and c) required studying serodiscordant couples but I don't think most studies of a) did.

This is a good point. We have clarified the text here.

The paragraphs starting at line 40 read like methods, not introduction

We think some of this material is useful to foreshadow what we're going to do in the body of the paper, but we have trimmed the technical details.

The term "epidemiological structures" is used throughout to mean the dynamics of the contact process (partnership dynamics), but this is not a common phrase and is not at all descriptive of what you want it to mean.

We have changed this term to "contact structure" throughout.

Methods: *If you refer to parameters here, they must be defined here, not just in the SI (i.e. for c_u , c_w , etc)*

We do give a full table of parameters in the main text (Table 1); however, we did not refer to it until the Latin hypercube sampling section — we now refer to it near the beginning of the Methods section.

What is the rationale behind starting at SPVL of 3? All of your results about the general time course of the mean SPVL depend on this. If it is simply to compare to Shirreff, then should state this.

Thanks to this comment, we realized that we did not say enough in the main text about our choice of initial conditions. We have added a subsection in the Methods section about initial conditions; we have also extended our discussion of our sensitivity tests of initial conditions (emphasizing that the peak SPVL is relatively insensitive to starting conditions).

Figure 1 is not well explained in the text or the caption. What is the point of these comparisons? Like where do these three values of r come from? Which one did Shirreff use? Same for other quantities.

We have added some text here.

You can't assume that your reader knows exactly what was done in Shirreff's paper. Instead of always saying what you did relative to Shirreff, the comparison between your methods and theirs should be written in a way that is clear, even if the reader has never read Shirreff.

We have added detail; we believe the presentation of our methods can stand on its own (although we still relegate

minor technical details to the supplementary materials, so as not to overwhelm the reader). We have also changed presentation more generally so that we don't lead by saying "as in Shirreff et al ...", which might make readers think that they need to know Shirreff's material, although we still try to be clear where we are following Shirreff et al/Champredon et al and where we are breaking new ground.

The description of the different contact dynamics is confusing. Again, don't assume readers have read Shirreff or Champredon - clearly explain what the models are. What you name the variables (e.g. SI, II) is not important, since you don't list the equations in this section anyways. The paragraph starting at 167 is more clear, but makes no sense to have this AFTER the more detailed description of the differences between the models - should be first. Start simple, add details.

We have moved the later paragraph up as suggested, and have rewritten the relevant paragraphs to clarify them.

Why did you assume a constant population size - that is, replace any dead individuals immediately with new susceptible ones? In this case, the limitation on susceptible that occurs near epidemic peak is greatly dampened, and is definitely going to change the infection dynamics and probably also the virulence trajectory. Since the goal of the paper is to add more realism by including more realistic partnership dynamics, then this seems like any easy place to avoid removing realism.

We made this assumption following Shirreff et al, for simplicity. We now describe the SIS assumption explicitly in the main text; we also constructed additional models with vital dynamics, and report on how this assumption changes our results (not much).

Results: *I would highly recommend that the figures and descriptions of the results refer mainly to SPVL, and augment with results on time to progression. It is just much simpler and more intuitive to think about SPVL. SPVL is the only quantity that is actually observed in most cases. Referring to time to progression is very confusing because LOWER time to progress = HIGHER virulence so just takes more effort to interpret.*

We agree that this causes some difficulties. Earlier drafts actually focused on SPVL as the output measure, but we thought that non-technical readers would have a better feeling for the real-world significance of progression time; furthermore, because progression time (unlike $\log(\text{SPVL})$ is on a ratio scale, it becomes reasonable to quote results in terms of proportional changes. We are taking the reviewer's advice and switching back to SPVL.

Throughout the paper you refer to the conclusion that the simplest and most complex models give the same results, and that models of intermediate complexity give different results. But none of the figures are really presenting this point, as someone looking at the figure probably does not remember the details of each model based only on its code-word, and the series in the figures do not seem to be organized in terms of model complexity.

This is a great idea; we have reordered the model presentation in the figures so they do reflect a complexity series (i.e., starting with "random" and ending with "hetero").

First paragraph: Don't understand "As r decreases from 0.084 to 0.42". Where did these r values come from? Is r changing over time? Or over different iterations? Why not just fit to the observed r value? And this change you've described is an INCREASE, not a decrease.

We have changed the wording, and corrected a typo (the second value should have been 0.042 — sorry!) As we say in our responses above, we have added more information about the choice of r , $I(0)$, $\alpha(0)$ in our sensitivity analyses (Figure 1).

Why are you changing the initial infectious density?

As part of a sensitivity analysis, to determine the effect of $I(0)$ on the results (since we see that the results are relatively insensitive, we fix a value of $I(0)$ thereafter)

Line 306: Why would the speed of virulence evolution affect the maximum attained virulence? Doesn't make sense. Should just influence the time to peak.

Because the total evolutionary change in SPVL over the course of the growth phase is equal to the time until the epidemic peak (years) times the speed of virulence evolution (change in SPVL/year); the faster evolution proceeds, the closer SPVL can get to its growth-phase optimum (i.e., to the SPVL that maximizes r).

It is very hard as a reader to maintain any interest in the results of different models when there was so little description of them, and, there is just too much detail for one to store in one's head throughout the paper. I would recommend having a diagram, or a table, that clearly compares each model, for what features it does and doesn't have, so the reader can keep referring back to this to try to make sense of the findings.

This is a great idea. We have added a figure that shows four diagrams (one for each of the central models), with a colour scheme matching the colours used to denote models in later figures.

I don't understand why Figure 5 shows the unscaled parameters, when in reality you had to scale the parameters. Does not make sense to show summary statistics for scenarios that you know are unrealistic (because they would lead to epidemics with r values that do not match the observed value)

For clarity, we have changed Figure 5 to show the scaled parameters.

Figure captions are not publication level. Some general advice on figure captions, taken from another source: Figures should be able to be interpreted relatively independently of the text. Most readers will go straight to the figures before reading the text in detail, and so the captions should not assume the reader has internalized every detail of the text preceding the figure citation. Figure captions should walk the reader through what experiment (simulation) was done, what analysis (statistics) was done on the data, and what quantities are shown in each axis. 100-200 words is a normal length. The first sentence (might be called the title), is often bolded, and can be either a summary of what data (or model, etc) the figure is showing, or a statement of the main results. Try to be consistent. In general, figures are supposed to describe results, not conclusions. Next, describe each panel of the figure, referring to it by label (eg A)). For each panel, or at the end of the caption (if this info is the same for all panels), describe the values of any important quantities that were held constant in during the experiment that generated the figure, state any important assumptions, and state sampling size or time if relevant. Avoid using parameter symbols or abbreviations, but if they are used, provide definitions in the caption. These conventions are often ignored in technical mathematical biology papers (i.e. those in JTB), but are crucial for making theoretical papers palatable for interdisciplinary audiences.

We have thoroughly revised the captions.

Discussion: *Would be helpful to start off by summarizing the main goals of the paper and then the main findings of the paper. Then, can get into limitations* We have added several paragraphs at the beginning of the discussion.

It would be very helpful to be able to give a intuitive or at least qualitative explanation for why these mixing patterns have such an influence on virulence. We discussed these issues some in the paragraph starting “The most striking aspect of the univariate comparisons ...” (p. 15 of previous ms., p. 18 of current version). The second paragraph of the Discussion now goes into more depth.

Reviewer #2: *The authors have appropriately dealt with my previous comments. Furthermore, I think that the addition of a model with heterogeneity in sexual activity is useful.*

Thank you!

Reviewer #3: *p. 10 if not counteracted by increased mortality among hosts infected by these reproductive strains.*

We have clarified this sentence by explicitly framing “reproduce more” as a net increase in lifetime fitness, i.e. increasing \mathcal{R}_0 .

p. 11 instead of calling this “epidemiological structure” it may be more appropriate to refer to it as “contact structure”. The authors investigated the effect of adding structure to the pattern of sexual contacts. There is nothing intrinsically “epidemiological” about forming (stable and/or once-off) sexual relationships.

We have changed this wording to “contact structure” throughout

p. 11 Given the long time horizons of the simulations (500-4000 years), this simplifying assumption [SIS model] may introduce bias, but unfortunately the magnitude and importance thereof is unknown. In more realistic models, one would expect that the high-risk susceptible individuals get depleted at a higher rate than the low risk susceptibles, with the epidemiological effects of this risk gradient playing out over a small number of decades. If or how this dynamic affects the evolution of virulence is unclear but if the goal is explicitly to study this evolution, it is problematic that all models under study are unable to take these demographic effects into account, by design.

We have implemented vital dynamics in our models and briefly discuss the impacts of this assumption in the Results. Overall, implementing vital dynamics increases the peak SPVL and mean SPVL (both by about 0.1 log₁₀-SPVL units). Time to peak remains similar, although it increases from about 250 to 300 years for the “instswitch” model.

p. 12 As noted by reviewer 1, the decision to ignore temporal variation in infectiousness over the main stages of HIV infection is a serious limitation of this study. In particular, the epidemiological contribution of the spike in HIV viral load during acute and early HIV infection depends on the assumed contact pattern. In models where individuals are not “trapped” in serially monogamous relationships, recently infected individuals can make efficient use of this phase as they can transmit to new or additional sexual partners as soon as they are infected. In contrast, in models that only allow for stable partnerships, this spike in infectiousness will “go to waste” as the newly infected individuals are most likely still in the relationship with the partner that they acquired the infection from during these initial few months after HIV acquisition. For highly virulent viral strains (high SPVL and short survival) such window of opportunity is more important than for viral strains that allow the host to transmit for a decade or more.

We now discuss this issue (l. 500).

p. 12 Showing that the model with only pair formation dynamics (Shirreff et al.) produces qualitatively similar results to those from the Shirreff model, does not prove or suggest that the effect of adding additional extra-couple transmission is robust against the model’s structure with respect to the stages of HIV infection and associated infectiousness levels.

This is essentially the same point as above; we have tried to temper our conclusions on lines 60 and 500.

p. 21 The authors touch on two weaknesses of their analysis here: 1. The level of sufficient realism is questionable, despite great efforts to take into account various sources of variation (with several thousands of compartments as a result). 2. The real-life implications of their main metrics (time to reach peak virulence, time to reach equilibrium virulence, progression time to AIDS at peak virulence) is unclear, because in all model scenarios (strain-specific) transmission and disease progression/mortality parameters stay constant over time. However, dramatic increases in the uptake of HIV treatment (ART) have been observed across the world, and trends in sexual risk behaviour may have changed both in the pre-ART and ART era.

We agree that our analysis is not by itself enough to assess the eco-evolutionary dynamics of HIV in a setting that is realistic enough to apply in a public policy context. However, we still feel it represents a valuable contribution to the HIV literature, by providing guidance to other modelers on both how and why eco-evolutionary conclusions can depend on model structure.

p. 22 The model with heterogeneous levels of sexual activity (the “heterogeneity” model) is an extension of the pair formation + epc model. However, in principle, this extra layer of heterogeneity in sexual contact rate could have been added to the other 5 “base” models as well. In the current comparison it is not possible to assess the relative effect of adding extra-couple and uncouple transmission versus contact heterogeneity.

We extended our analysis to apply both heterogeneity and vital dynamics (separately) to all four of the base models (pairform/instswitch with and without extra-pair contact); we report our findings in the last paragraph of the Results section and in Figure S4 (Supplementary material 2).

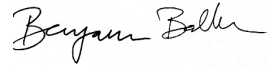
p. 23 This sentence [“These differences are practically as well as scientifically important”] seems to belong better in the discussion section. Here, the authors do not explain what the practical or scientific importance of the differences is.

We removed this sentence.

p. 27 The main finding of this analysis seems to be that allowing “unstructured” transmission events, be it extra-couple or uncoupled, influences the expected trajectory of SPVL evolution over rather long time scales. As Reviewer 1 notes, this is indeed a rather theoretical finding. In my option, two key questions remain unaddressed: 1. Why should we care more about the effect of this aspect of the assumed contact pattern than the effects of other aspects of the contact pattern and disease process (e.g. age-mixing, stage-varying infectiousness, ART status, depletion of high-risk susceptibles)? 2. What are the practical implications of these findings for HIV epidemiology? For HIV prevention and treatment? Does it help us in any way to interpret empirical time-trends in SPVL or disease progression?

We agree with the reviewers that this is a “rather theoretical” finding, but we think it is important and useful nonetheless. As we mention in the discussion, the decision to examine the effects of contact structure does not mean we can feel free to ignore all other aspects of epidemiological realism; on the other hand, data for parameterization and time for building and checking models is limited, so we have to prioritize in some way. We would argue that both quantifying the effects of contact structure and understanding how they work is a useful, albeit incremental, step toward reliable models of HIV eco-evolutionary dynamic.

Sincerely,

A handwritten signature in black ink, reading "Benjamin Bolker". The signature is written in a cursive, flowing style.

Benjamin Bolker