Dear Editor:

Thank you for the chance to revise this submission. Below please find our responses from the previous round.

Reviewer 1

This revision is substantial, being approximately four pages longer than the original, and with seventeen more references. The title has changed on the manuscript but remains the same on the cover sheet. The authors should standardise on 'reproduction number', which is correct, instead of 'reproductive number', or at least be consistent. See Figure 1. They should also use 'infection' rather than 'disease' in many places.

We have changed 'reproductive number' to 'reproduction number' and replaced 'disease' with 'infection' in appropriate places.

Table 1. Suggested additional reference Roberts & Nishiura 2011. PLoS One 6: e17835.

We have added the reference to the table.

Table 1. typo row 8 column 3.

We have fixed the typo.

Page 6 "taking only non-negative values makes it more biologically realistic than the normal". What does this mean? The normal only takes positive values too.

We mean that the domain of a normal distribution extends from negative infinity to positive infinity. We have made this clearer in the current text: "restricting the domain to only non-negative values makes it more biologically realistic than the normal."

Page 6 "its theoretical properties and practical importance have not yet been explored in depth." This is nonsense. There are whole books about the gamma

function and its properties are taught to undergraduates.

We have changed the text to clarify this point: "its theoretical and practical importance in explaining the r–$\mathcal{R}$ relationship has not yet been explored in depth"

Page 8, line 4. "converges to exp"

We have followed the reviewer's comment in changing the text.

Page 8, paragraph "Characterizing . . . higher R" could be a lot clearer. So could the next paragraph.

We have followed the reviewer's comment to make the paragraphs clearer:

" Characterizing the r–$\mathcal{R}$ relationship with mean and coefficient of variation also helps explain results based on compartmental models, because the mean and variance of the generation interval is linked to the mean and variance of latent and infectious periods. For example, less-variable latent periods result in less-variable generation intervals, whereas less-variable infectious periods result in both less-variable and shorter generation intervals (Svensson *et al.*, 2007). This explains the apparent anomaly between earlier results: when mean generation interval is held constant, less-variable infectious periods only reduces variation in generation intervals, leading to lower $\mathcal{R}$; when mean infectious period is held constant, less-variable infectious periods reduces the overall length of generation intervals, leading to higher $\mathcal{R}$ (Wearing *et al.*, 2005).

There is a simple intuition for this result. Initial exponential growth of an epidemic is largely driven by shorter infections. Increasing variation in latent periods results in increase in number of infections with early progression and a faster epidemic – equivalently, lower $\mathcal{R}$ is required to match the desired value of r. Increasing variation in infectious period results in increase in number of infections with early recovery and a slower epidemic – equivalently, higher $\mathcal{R}$ is required to match the desired value of r "

Page 9 typos. 'generating a realistic', 'data is are limited.', 'latent period to a an'

We have followed the reviewer's comment in changing the text.

Page 10 Suggest "We investigated this our approximation approach using three different examples."

Figure 3. What is the justification for including ρ ¿ 1?

While $\rho < 1$ is the relevant range for the Ebola outbreaks, we chose higher ρ to illustrate that the gamma approximation works over a broad range of growth rates in this example.

Page 13 "there is small little difference"

We have followed the reviewer's comment in changing the text.

Page 14 'for rabies' not 'for the Rabies', gamma lower case

We have followed the reviewer's comment in changing the text.

Page 15, first paragraph of discussion change to present tense.

We have followed the reviewer's comment in changing the text.

Claim the gamma approximation was introduced in [29] is incorrect as [37] predates it.

[37] approximates latent and infectious period with gamma distributions; this is different from approximating genertaion-interval distribution with a gamma distribution.

The bibliography style should be consistent. A lot of work is needed here

S1.1 'For this model, the generation interval distribution', 'the squared coefficient of variation', 'relationship under for the SEIR'

We have followed the reviewer's comment in changing the text.

Reviewer 2

The revised manuscript addresses some of the issues that were raised.

For the original paper, I thought that the main novelty was the proposal of an estimator of R based on the gamma distribution. This was incorrect, there were several earlier publications that have introduced this paper. The authors have now included a literature review and a useful table (Table 1) to guide the reader through what is known about the topic.

The authors do not provide any guidance on the generality, other than stating the approach works well for the examples that are studied. This is made clear in the revisions.

We appreciate your comments. We tried to make the contribution of our paper clear in our previous revision.

Reviewer 3

In this paper authors explore the relationship between R and the intrinsic growth rate "r" under the premises of exponential growth of epidemics with examples for Ebola, rabies, and measles. While the theoretical parts of the paper are sound, the practical parts are not well connected to empirical observations. In particular, the behavior of an outbreak during the initial phase is an approximation based on a list of assumptions and conditions.

Notation: Authors denote the basic reproduction number as R while it should be denoted as $\mathcal{R}_0$.

We have followed the reviewer's comment in changing the text.

It should be noted that R0 only applies at time 0 when the epidemic takes off.

We made this clearer by changing the first sentence of the first paragraph: "Infectious disease research often focuses on estimating the basic reproduc-

tion number $\mathcal{R}_0$, i.e., the number of new infections caused on average by a single primary infection in a fully susceptible population."

Assuming exponential growth is a strong assumption as it implies that the growth rate r is constant in order to maintain an unchanged R0 over generations during the initial growth phase (all individuals carry the same intrinsic growth rate and r is independent of time). Yet, the ascending growth phase of many outbreaks for different infectious diseases follow sub-exponential growth rather than exponential growth as recently noted in several publications (see e.g., Viboud et al. Epidemics 2016; Chowell et al. J Roy Soc Interface 2016).

We are aware of sub-exponential growth of epidemics. We did not account for sub-exponential growth models in this manuscript to provide a clear, concise message. The goal of this paper is not to advocate using exponential assumptions but to advocate biological intution behind the role of generation intervals in relating the growth rate with the reproduction number of an epidemic. Qualitative relationships are expected to be robust whether an epidemic is growing exponentially or sub-exponentially.

After time 0, any systematic changes in the growth rate r will affect the effective reproduction number R(t), and R0 no longer applies. Hence, assuming that epidemics grow with a sustained R0 beyond a couple of generation intervals may be unrealistic. That is, one has to keep in mind that exponential growth may only be a good approximation for the very few generations under strong assumption that R0 is the same for all individuals including the individuals seeded at generation 0.

JD: can you answer this?