This manuscript brings a timely contribution to the current COVID-19 pandemic by studying the confidence and bias of the infected population testing and disease spread. The major contribution is the authors developed the statistical compartmental models to assess the bias and accordingly simulate the dynamics. This manuscript would interest to general readers of PeerJ, I therefore think it could be acceptable for publication after introducing more real-world experiments, as well as addressing the following problems.

1. Line 80. population rBT + (1-r)T, this formula is bigger than the total population T, please clarify.
2. Line 95, Eq. 9 and 10. λ- and λ+ need their denotations. Maybe they are the expectations of the number of positive and negative tests, but the mathematical expectation of a Poisson process is λ parameter, which is confused with λ(t) and may lead to misunderstanding. I therefore suggest the authors replace λ(t) by a new symbol.
3. Line 81, Line 95 and Figure 3. What is the exact mathematical or epidemical meaning of B? is B(t) the time-dependent B? Fig. 3 should specify the variance of B. Also I cannot find why 1000 unbiased samples are reasonable. Based on fig.3, if aiming to find 90% of the cases with 90% confidence, using more than 1400 unbiased and 1200 biased samplings is more justified.
4. Line 190. Section “Trends in COVID-19 data”. This section seems to have little relationship with your proposed methodology. More detailed cases are needed to present the feasibility of your model.
5. Fig. 5 and 6 should better use the same legend.
6. Appendix A. SAIR model is easy to follow, but it is unclear that these betas have the similar strategies with Eq. 14. Using parameterization is a good idea, but a brief introduce on Eq. 14 is important, for instance, Kuznetsov et al. (Bifurcation analysis of periodic SEIR and SIR epidemic models, 1994) parameterized beta, which is seasonally varying.
7. Appendix B, Line 340. What the simplifying assumption is.

8. Appendix B, Eq. 16. The first “=” should be “≈” because this formula is true If and only if the time interval is infinitesimal.