26 February 2021

Dear editors,

Thank you for sending reviewer comments on our manuscript. These have provided important feedback and been immensely useful in improving the paper.

We attach a revised version of the manuscript which addresses the comments and suggested changes. Below we have summarised reviewer comments in italics, together with our responses structured by section in the paper.

Thank you again for your patience in editing the manuscript.

Best wishes,

Katharine Sherratt

**General**

*Reword title*

* We have edited the title to better reflect our focus on how surveillance data biases can provide information on subpopulation transmission.

*Formatting*

* We have corrected *Rt* typesetting
* We have capitalised “Table” and “Figure” references
* We have removed the use “*Rt*s” throughoutand replaced with *Rt* estimates and similar language – thank you for suggesting this
* We have replaced a mistaken reference to Fig SI1A with the correct Fig SI2A.

**Methods**

*Test out UK specific delays if time permits.*

* Unfortunately, this was not completed in time. We are exploring using data from the CO-CIN observational study to do this in future work.

*Clarify delays for hospital admissions and test positives are treated as having the same delay from onset (and therefore the same lag from infection to observation).*

* We agree this should be made clear. We have included this explicitly: see Methods, paragraph 7.

*Page 5 – extra explanation as to how the uncertain distributions were then sampled. For example, how are the uncertainty in both mean and standard deviation captured. When estimating these delays, mean and standard deviation are coupled, so is the uncertainty generated from e.g. a posterior sample of mean and standard deviation pairs, or are means and standard deviations sampled assuming the uncertainty is independent?*

* As the reviewer notes when estimating these delay distributions, the mean and standard deviation are coupled. As our method uses two modelling steps, with the first estimating the delay distribution parameters and the second using summarised posterior estimates (in order to reduce computation) as priors, we assumed independence – as noted by the reviewer. This may have increased the uncertainty of our estimates, though we note that as these distributions are refit in the second modelling stage (though with little data to inform them) this issue may have been mitigated.
* We have expanded on this approach in the methods: see Methods, paragraph 6.
* We thank the reviewer for highlighting this as jointly fitting these distributions along with reconstructing infections is an interesting area for further research.

*P5, L37: Is this prior informed by the data not equivalent to “using the data twice”?*

* Yes, we use a small subset of the data in both the prior and model fitting. Although a limitation, we believe this will only have affected the first few estimates, and that this was the only viable modelling option. Alternative methods struggle with model identifiability. We also note this assumes that initial observed growth is equivalent to unobserved growth, which we believe is justified and is the default for other approaches that do not model latent infections.
* We have clarified this in the text: see Methods, paragraph 3.

*P5, L38: How are imputations done?*

* Imputations are done in the model, with the initial number of infections as described then weighted by generation time and *Rt*.
* We have clarified this in the text. We have also updated the wording to reflect that imputation is generally used to describe modelling missing data which is not the case here: see Methods, paragraph 3.

*What is an uncertain generation time p.m.f? p.m.f. is uncertain, or that the p.m.f. is known, but models a stochastic outcome?*

* The generation time distribution was assumed to be known and fixed over time for each MCMC step, with a prior placed on the mean and standard deviation. This was a pragmatic choice to maximise uncertainty: for example, sampling at each time step would potentially narrow CIs if there was little data to inform the posterior of the generation time.
* We have edited the text to include this: see Methods, paragraph 3.
* We are exploring alternative methods for estimating the generation time distribution and for modelling it jointly along with the other distributions currently defined using literature based priors.

*Mathematical details of the modelling that was used, rather than simply references to other papers – e.g. the full Bayesian model specification (with priors)*

* We have added model specifications and priors: see Methods, paragraph 4.
* We have included the model specification in math notation.

**Results**

*Page 6 line 55 – “However, as much as spatial variation, the data sources used to estimate Rt influenced the earliest date of epidemic decline.” – edit for clarity*

* We agree the sentence was unclear and we have edited this: see Results, paragraph 2.

**Discussion**

*UK-specific vs global delays - would this improve/worsen the discrepancy between admissions and deaths*

* We used public data wherever possible and compared this to available confidential data in the UK. This led to using global delays for admissions/cases and UK specific delays for deaths. The discrepancy should not have been substantially impacted, but using the public linelist data may have increased uncertainty in *Rt* estimates.
* We have added substantially more detail and our justifications: see Discussion, paragraph 9.

*Implications of delays being the same for cases / hospitalisations - for example, with the higher testing rate rolled out over summer and wider community testing, the delay from symptom onset to testing might have decreased, whereas the delay from onset to hospital admission will not have experienced the same change.*

* Using the same delays could explain some of the variability in *Rt* estimates from test-positive case data compared to hospital admissions. We had no data over time on delays from symptom onset to reporting in each data source to test this; however, we have mitigated some of the likely impact by using independent sampling over an uncertain delay distribution for each set of estimates.
* We have substantially expanded upon this point in the text: see Discussion, paragraph 8.

*Page 9 line 21 - local nosocomial outbreaks could have also contributed to this discrepancy*

* Thank you for suggesting this, which we agree this is a strong explanation.
* We have included an explanation of how a discrepancy in *Rt* would be caused by biased detection of nosocomial outbreaks in testing data combined with the healthy-worker effect.
* We have also re-arranged the discussion to link together this range of explanations associated with uneven sampling across the age-severity gradient over time: see Discussion, paragraphs 4-6 and specifically paragraph 4 for nosocomial infections.

*Discuss whether pooling estimates might help provide a more robust estimate of Rt, or whether it’s better to present multiple estimates to policy makers*

* Thank you for raising this, which is an important addition to the discussion.
* To summarise, we recommend against pooling estimates. It is unclear how weights would be assigned based on likelihood to estimates from different data sources. We also believe that when sampling biases are understood, the variation in concurrent *Rt* estimates can be used to gather more information about population transmission than any single estimate.
* We have added these points and brought them out more clearly as recommendations: see Discussion, paragraph 14.