Response to comments from RSS meeting

We thank all the discussants for their excellent and constructive remarks. We generally agree with all of them, and take this opportunity to integrate some of these comments with further considerations.

Different information provided by R and r

One of the cornerstones of our manuscript was the comparison between the reproduction number R and the growth rate r, including the different information they provide. Although discussion points raised in previous sessions aimed to conclude that one concept is uniformly "better" than the other, our position is much more in line with the spirit of Kucharski and Scalia-Tomba's comments (also voiced by many other discussants throughout the meeting) concerning the fact that different tools might be more or less appropriate depending on the question at hand or the purpose of the analysis. For example – as we mention at the end of the manuscript – if all we want to do is to track whether the epidemic is growing or declining or how the previous change in policy has affected the epidemic trend, then R provides no more information than r, might be more sluggish at responding to changes in trends, and requires increased model complexity. Similarly, if what we need is to provide a short-term projection of the current trend under the assumption that conditions are not changing, "interpretability" (Kucharski) or "knowledge" (Scalia-Tomba) might not be the primary concern. In this case, again r requires fewer modelling assumptions and fewer model ingredients (such as the generation time distribution), which themselves come with associated uncertainty, and therefore might be preferable to R. If, however, separation of the various mechanisms involved in the transmission process, the likely impact of future control options, or considerations about critical immunity thresholds are the primary interest, then R might be more informative.

In [1] and in a comment in Session 1 of this meeting, we argued that robust estimates of the unconstrained growth rate r_0 are essential in the early epidemic phase. Especially in combination with estimates of delays between interventions and the observation of their effects, they provide crucial guidance about the timing of initial interventions, which the time-insensitive R_0 cannot provide. We do not expand further on these important points here, as they are extensively covered in [1] and Session 1. However, it is particularly relevant to this discussion to note how Figure 4 in [1] displays a relatively narrow range in the early (until mid-March 2020) estimates of R_0 despite the corresponding r_0 estimates being more variable, likely due to the limited knowledge on the generation time distribution at the start of the pandemic. The inconsistency between the range of R_0 and r_0 is a natural result of different modelling assumptions, but might be of particular concern if resulting from a propensity of research groups providing new estimates of R_0 to remain in line with previous, highly cited, values. This case may also exemplify a potential tendency to avoid uncertainty when communicating results to the public (Scalia-Tomba).

Another valuable difference between the two concepts that is worth clarifying, given its recent relevance, concerns the emergence of COVID-19 variants and potential assessment of their transmission advantage and increased severity [2,3,4]. Growth rates can describe the speed at which new infections occur in a population, or could simply be used to describe the speed at which any observed data stream of interest is growing: for example they can be used to describe the growth in hospital admissions, or the growth in positive cases, while remaining agnostic about the source of such cases. Transforming a growth rate into a reproduction number instead requires the assumption that those cases are the exclusive result of transmission in the population of interest. For example, the UK has witnessed a distinct growth in positive cases of the Delta variant since April 2021, fuelled initially by importations, primarily from India. With limited or delayed information about the incidence of cases among travellers, the temptation to blindly apply methods that directly estimate R from observed cases would result in incorrect estimates of the transmission advantage of the Delta variant. More refined methods, which separate importations from community transmission (and hence require relevant data), would reach more robust conclusions.

We note, however, that there are many cases during outbreaks where the value of R is actually key in determining management. Consider, for example, the 2012-13 measles outbreak in Wales [5], during which various local authorities experienced periods of extremely short doubling times, often close to weekly. The main intervention against this outbreak was catch-up vaccination, but the levels of additional vaccine coverage required depends on R. This epidemic therefore provides a good example of the use of different quantities: the doubling time provided evidence for the urgency of fast action, while consideration of R informed the level of intervention required.

Modelling the observation process

A common theme across numerous comments concerned the modelling of the observation process linking the unobserved transmission dynamics to the data collected in practice, to estimate how the delays between infections and observations change over time with changing control policies, age distribution of infections, improvements in treatment, vaccine coverage etc. Undoubtedly this is easier said than done when tracking the epidemic in real-time, due to the large demand in terms of data and, importantly, of information on the data collection mechanisms and reliability. For example, answers to questions like “how has the data been collected?” or “have testing efforts been increased in a specific region due to an assessed or hypothesised local outbreak?” have been, in the authors’ experience, really hard to obtain, especially in real time. Models can capture apparent inconsistencies or suspected changes in the data streams, for example by allowing some parameters (e.g. the infection-fatality rate) to vary over time. However, proper assessment of the causes of observed changes or inconsistencies requires a lot more data and possibly expanding the model to introduce further mechanistic details to correct for biases and disentangle

transmission from other confounding processes. This assessment is likely only possible (if ever) retrospectively.

If sufficient data is available, Riley's comment opens up the question about whether assuming discontinuous changes in the transmission rate is appropriate, as opposed to assuming smoother transitions. We agree that the answer to this question might motivate slightly different modelling structures. However, it could be argued that, since incidence is never directly observed and what can be observed is effectively a convolution with some distribution (infection-to-testing or infection-to-hospitalisation), it is difficult to assess whether a smoother change in R (like those estimated with non-mechanistic models) or a step-wise one (like those used for simplicity in the mechanistic models discussed in Sections 3.1.2 and 3.1.3 of the manuscript) are more appropriate. Ultimately, they both appear to fit the same "smeared" data equally satisfactorily.

Challenges in the estimation of population immunity

As the interest shifts from simply fitting epidemic trends or generating short-term projections assuming stable conditions to disentangling the effect of various factors affecting the infection and observation processes, increasingly complex mechanistic models are needed. As noted in Diggle's comment, one of the most difficult factors to properly assess is the fraction of the susceptible population. There are multiple reasons for this, including: large scale survey studies (e.g. ONS and REACT in the UK) typically allow estimation of current prevalence in the community, but this is not the same as cumulative incidence and in any case such studies are not available from the beginning of the pandemic and do not necessarily cover specific subpopulations particularly at risk (care homes, health care workers, etc.); information on seroprevalence does not translate directly to immunity; vaccine effectiveness against infection is difficult to estimate precisely and changes over time with an evolving virus; and loss of immunity due to infection or vaccination is similarly difficult to quantify. The lack of precise estimates of susceptibility in the population means prediction of the timing and height of the epidemic peak, which are primarily driven by the depletion of susceptibles, is fraught with difficulties. Correctly predicting the peak is further complicated by the intrinsic problems of assessing the right level of individual heterogeneity to capture in the model structure and the possible lack of suitable information to parameterise it. Finally, it is often not possible to predict changes in individuals' behaviour and control policies between the time of the forecast and the peak.

The difficulty in estimating the fraction of the susceptible population also affects the accurate estimation of R_c, and hence the impact of interventions. However, similar sets of interventions applied in different phases of the epidemic would in theory have the same R_c but different R_e and could provide information on the reduction in the susceptible population that occurred between such phases. The presence of different variants and natural changes in individuals' behaviour (e.g. due to intervention fatigue), may

however falsify the assumption that R_c remained the same even for the same set of interventions. When different combinations of interventions are implemented at different times, the situation is even more complex. Scalia-Tomba suggested potential extensions of the concept of R_c (one value accounting for all forms of control policy altering natural contact patterns) to provide different reproduction numbers associated with each separate control policy. However, as we explain in the manuscript, the compounded effect of control policies is not necessarily the sum of the effects of each single control policy. To draw a parallel with network theory, a static network might remain connected (each node can be reached by any other node) if each of two distinct sets of links were to be removed separately, while removal of both sets simultaneously might break the network apart and render it disconnected, thus qualitatively changing the spread of an infection on it. However, crude approximations that attribute an "R budget" to each control policy have been used in practice to help discussions in the UK response. Effort to formalise these ideas and clarify when they offer a useful approximation and when they do not is undoubtedly valuable.

Kucharski states that "vaccination decouples the relationship between available datasets", and we agree that, as vaccination is rolled out, the link between infections, hospitalisations and deaths will change. However, we deem expressions like "dilute" to be more appropriate than stronger ones like "decouple" or "break" in terms of communication to non-experts and expectation management: if conditions remain constant over time, in theory the same exponential growth should be observed in all these data streams, though with smaller proportions of cases resulting in deaths and hospitalisations. Hence, a large number of cases could still lead to substantial hospitalisations and deaths.

Panovska-Griffith echoes Kucharski in pointing out how the profound change due to mass vaccination poses the question of how to combine different data streams in order to monitor the epidemic status and inform interventions. A natural follow-up question is whether R remains a useful quantity to be monitored, or whether alternatives such as the growth rate in the number of hospitalised cases should be preferred. Our opinion is that estimates of R and r should not be abandoned for several reasons. First, because estimates of R remain relevant as indicators of ongoing transmission, allowing an understanding of the new relationships between the data streams (e.g. new estimates of the infection-fatality rate) and providing an alert of impending changes in severity. Second, comparison of the growth rates of infections and hospitalisations is important to investigate the reasons for possible discrepancies, since the two data streams should in theory have the same growth. Finally, estimation of R and r would not require additional efforts, as they have been regularly monitored since the beginning of the epidemic.

All these considerations have hopefully further highlighted the range of subtleties that accompany the concepts of reproduction numbers and growth rates. This reinforces the

case for effective communication of such subtleties, and demands that continuous effort should be devoted to their clarification and to education of policy makers and the general public. We are therefore very grateful to the organisers of this meeting, the other authors and the discussants, for their valuable contributions to this cause.

References:

[1] Pellis, L., Scarabel, F., Stage, H. B., Overton, C. E., Chappell, L. H., Fearon, E., ... & University of Manchester COVID-19 Modelling Group. (2021). Challenges in control of Covid-19: short doubling time and long delay to effect of interventions. Philosophical Transactions of the Royal Society B, 376(1829), 20200264. https://royalsocietypublishing.org/doi/full/10.1098/rstb.2020.0264

[2] Davies, N. G., Abbott, S., Barnard, R. C., Jarvis, C. I., Kucharski, A. J., Munday, J. D., ... & Edmunds, W. J. (2021). Estimated transmissibility and impact of SARS-CoV-2 lineage B. 1.1. 7 in England. Science, 372(6538). https://science.sciencemag.org/content/372/6538/eabg3055.abstract

[3] Challen, R., Brooks-Pollock, E., Read, J. M., Dyson, L., Tsaneva-Atanasova, K., & Danon, L. (2021). Risk of mortality in patients infected with SARS-CoV-2 variant of concern 202012/1: matched cohort study. BMJ, 372. https://www.bmj.com/content/372/bmj.n579

[4] Nyberg, T., Twohig, K. A., Harris, R. J., Seaman, S. R., Flannagan, J., Allen, H., ... & Presanis, A. M. (2021). Risk of hospital admission for patients with SARS-CoV-2 variant B. 1.1. 7: cohort analysis. bmj, 373. https://www.bmj.com/content/373/bmj.n1412.short

[5] http://www.wales.nhs.uk/sitesplus/888/page/66389