

1 We thank the Royal Statistical Society for the opportunity to appraise comments from discus-
2 sants and the discussants themselves for providing feedback. As some of the comments cover
3 similar topics, we respond to the points raised rather than individual responses.

4 **Estimating accurate case counts (Diggle)** Entire editions of academic journals are dedicated
5 to infectious disease modelling efforts while proper use of data to inform the modelling has been
6 emphasized only recently (e.g. Held et al., 2020). The importance of data deserves highlighting
7 and it is noteworthy that one of the most detailed and often analysed datasets in the field dates back
8 to a measles outbreak in 1861 (recently re-analysed in Aaby et al., 2021). Without useful data,
9 we will not be able to estimate the susceptible and asymptomatic proportions of the population.

10 Strengthening and improving national and intergovernmental (coordinated by bodies such as
11 ECDC and WHO) disease surveillance and monitoring systems allows for improved early disease
12 outbreak detection. Such disease surveillance systems include both mandatory case reporting of
13 notifiable disease, sentinel surveillance systems, and also internet and news media, under the
14 umbrella of epidemic intelligence services. Disease surveillance requires certain amounts of
15 man power and resources to function and systems have seen increases in technological capacity
16 in recent years (Hulth et al., 2010; Groseclose and Buckeridge, 2017). Time series of infec-
17 tious disease cases typically arising from a surveillance system can easily be modelled using the
18 framework we used and presented. However, if the underlying data is flawed, so too will be the
19 outputs. We are cognisant of the adage “garbage in, garbage out”. While we are aware of many
20 funding opportunities for COVID-19 modelling, it is unclear how much emergency grant sup-
21 port has been given to strengthening current and future data gathering and storing infrastructure.
22 Utilising existing data mechanisms rather than “re-inventing the wheel” is paramount. Relat-
23 edly, there has recently been an attempt at rebranding the data-focused parts of infectious disease
24 surveillance as “outbreak analytics” (Polonsky et al., 2019).

25 In our own work examining the effect of travel restrictions to neighbouring regions on cases
26 in Switzerland we have recently considered both Italian and French case data (see Grimée et al.,

27 2021, for an initial analysis of some of the data) and have experienced two matters that caused
28 us to consider the data in further detail and not simply model it as-is. The first is that certain
29 case counts in Italian regions show changes from one day to the next which seem unrealistic. In
30 particular we have instances of zero (or even negative!) case counts followed by large counts.
31 The second is incoherence in case counts for French regions between data sets after changing
32 data providers. The Zurich case data does not suffer such problems, but certain cases may not be
33 captured by the surveillance system, and so there is a risk of underreporting.

34 **Underreporting (Diggle, Scalia-Tomba, and Kucharski)** Routine infectious disease surveil-
35 lance systems are prone to only capturing part of the disease prevalence and so provide an in-
36 complete picture of the burden. Specifically, not all infected persons will develop symptoms
37 (asymptomatic cases) and thus seek healthcare, whereby their case may not be reported in either
38 notifiable disease surveillance systems or sentinel and syndromic surveillance systems. The im-
39 pact of underreporting on endemic-epidemic models was examined by Bracher and Held (2020b)
40 and we are aware we need to correct for this in our larger Switzerland-wide analysis of school
41 closure, taking into account that underreporting may be age-dependent. The reporting also de-
42 pends on a correct clinical diagnosis (i.e. no misdiagnosis) and timely entry in the notification
43 system. Certain delays are inherent to the reporting system, e.g. the time between a test being
44 taken and sent to laboratory for analysis, and are usually corrected for using nowcasting (Höhle
45 and an der Heiden, 2014). Increased testing efforts are expected to change the reporting rate as
46 more asymptomatic cases will be captured.

47 **Metrics for communication between technical experts and policy makers (Scalia-Tomba,**
48 **Kucharski, and Panovska-Griffiths)** Our work is a “proof-of-concept” analysis and forms
49 the basis for an extended analysis of data from all of Switzerland and so the feedback will help
50 hone future efforts. Our paper provides expected case counts in order to investigate the effect of
51 school closures on disease incidence in the relevant age groups and shows that such an approach
52 works. Such case counts could be a metric reported in addition to the effective reproduction

53 number R and the growth rate r . For specific formulations of endemic-epidemic models, it is
54 even possible to estimate an effective reproduction number in addition to expected cases (see
55 Bracher and Held, 2020b, for details).

56 **Need for null hypotheses in infectious disease modelling (Riley)** We agree that there is a
57 need for well-specified null hypotheses to examine the effect of disease control interventions.
58 Null hypotheses may need to be born from benefit-harm assessments. The societal damage from
59 a public health emergency affects more than simple case counts. It is crucial to balance benefits
60 and harms, which policy makers do qualitatively, in a quantitative manner. As we are not in the
61 position to decide which measures to introduce or lift, we cannot determine with great certainty
62 what an “acceptable” number of additional expected cases is, but we like to stress the importance
63 of age in such considerations.

64 Related to this, we wish to briefly highlight an experience we have had during our work in the
65 previous year. To avoid creating unnecessary research waste and add to the gargantuan amount
66 of exploratory COVID-19 modelling papers, we submitted our work as registered research with
67 an associated study protocol (Chambers, 2019a,b). The preregistration was written according
68 to Van den Akker et al. (2020) specifications. One of the sticking points from our protocol is
69 how to determine a specific and testable hypothesis for our approach with associated rationale
70 (question 4 of the Van den Akker et al. (2020) specification). In the absence of well-defined null
71 hypotheses as requested by Riley, such protocols can be hard to complete.

72 Reviewers specialised in modelling analyses of infectious disease surveillance data do not
73 seem well-versed in the preregistered publication approach. The academic editor admitted to
74 finding reviewers with the required subject matter expertise who were also able to review pro-
75 posed procedures difficult. Finding reviewers for the myriad COVID-19 papers being released is
76 already taxing (Schwab and Held, 2020). It would seem following traditional publication meth-
77 ods (with review only occurring after the analysis is completed) are the ones used by the wider
78 field, albeit with pre-prints and providing access to a repository with their analysis code being

79 increasingly utilised (Brooks-Pollock et al., 2021). These approaches still do not allow the option
80 to appraise the methods before they are applied to data. Additionally, checks of data quality prior
81 to modelling (cf. the need for improved data) provide additional motivation for infectious disease
82 modellers to preregister their work.

83 **Comparing hypothetical control options (Kucharski)** While we used prediction retrospec-
84 tively, the model could also be used prospectively to predict the effects of a future control sce-
85 nario. The endemic-epidemic modelling framework is often used in probabilistic forecasting
86 (Bauer et al., 2016; Ray et al., 2017; Stojanović et al., 2019; Held and Meyer, 2020). Many
87 of the recent extensions to the framework consider aspects which need to be considered for
88 such forward-looking approaches (Held et al., 2017; Bracher and Held, 2020a) and incorporate
89 methodology used in weather forecasting. We have not personally examined future scenarios
90 of interventions using the modelling framework, as we have preferred to inform our work by
91 available data.

92 Informing the model with future hypothetical time-varying contact matrices would enable us
93 to examine the predicted number of cases under such hypothetical scenarios, e.g. returning to
94 baseline contact levels to represent fully reopening/removal of all social distancing measures. For
95 examples of how such hypothetical contact matrices may be constructed see Willem et al. (2020)
96 and Prem et al. (2020). Similarly to how we constructed our contact matrix with data on policy
97 interventions, Alleman et al. (2020) informed changes to a contact matrix with mobility data
98 and van Leeuwen et al. (2020) updated a contact matrix using time-use survey information. An
99 alternative would be to use contact surveys conducted during the COVID-19 pandemic (Jarvis
100 et al., 2020b,a; Feehan and Mahmud, 2020; Latsuzbaia et al., 2020). In the work presented
101 here—the pilot analysis of Zurich COVID-19 case data—we used a synthetic contact matrix
102 which is informed by demographic data as well as contact diary surveys (Mistry et al., 2020).
103 Demographic data has also been suggested as a way of “updating” older contact matrices for
104 newer use (Arregui et al., 2018) as the commonly used POLYMOD matrices are now some 16

105 years old and conducting a contact survey may be resource intensive.

106 **Changes in transmissibility and choice of age groups (Riley and Scalia-Tomba)** The con-
107 struction of our time-varying transmission weights is based upon informing a contact matrix by
108 policy indicators given as step functions. We have previously considered use of ramp functions
109 (as an alternative representation of changes in policy) in place of step functions. However, the
110 choice of slope in such a ramp function needs to be informed by relevant information. We have
111 not considered a smooth function as suggested by Riley. For simplicity, we continued our work
112 with the step function representation of policy (hence transmission opportunity) changes.

113 It is true that the construction of the time-varying contact matrices has assumed all members
114 of the population are in the same class with respect to factors that are not age. If information on
115 subclasses of interest (e.g. “responding”) is available to inform the model, it would be possible
116 to include an extended contact matrix including subclasses, meaning cases would also need to be
117 further divided depending on subclass status. If such a status is true for certain age groups, e.g.
118 the younger three, it may be better represented as a covariate with the same matrix structure as the
119 observed counts rather than increasing the dimension of the matrix to reflect the increased number
120 of classes. The goal is to include enough nuance that the transmission matrix is informative for the
121 groups of interest included in the model, but doesn’t incorporate unnecessary distinctions which
122 could mean artificially low disease counts would enter the model, and could cause convergence
123 problems.

124 For example, in our work we have not stratified cases by sex, as the patterns of case counts are
125 similar for each sex. It bears mentioning that summing the results from a multivariate endemic-
126 epidemic model may not yield the same as those found in the univariate version, as the interplay
127 between groups will not have been incorporated. A related issue is how sensitive the results are
128 to the choice of the age groups. We have tried to define the age groups in a reasonable manner
129 (school children, working adults, elderly, etc.) though it would be interesting to investigate how
130 sensitive the results are to other stratifications.

131 **Generalisability and vaccines (Kucharski)** While our modelling approach can easily be ap-
132 plied to other countries, when working with COVID-19 data for multiple regions, it is pertinent
133 that users of data gathered consider whether the case definitions and testing strategies are the
134 same across regions. If data is not harmonised in such a manner, conclusions may not be straight-
135 forward in multi-region comparisons. With regards to the roll out of COVID-19 vaccines, it is
136 important to know not just how many doses of vaccine have been given but also which ones
137 they are. To continue with the examples of the two countries considered, at the time of writing
138 (July 2021), Switzerland has only licensed messenger RNA vaccines (Comirnaty and Spikevax)
139 for use against COVID-19, while other options exist (e.g. adenovirus-based Vaxzevria) in the
140 United Kingdom, a nuance which might not be evident from numbers of proportion vaccinated
141 in each country. Furthermore the immunisation regimes are different, many younger Swiss resi-
142 dents are currently fully vaccinated with 4-6 weeks between shots while their British equivalents
143 are waiting up to 12 weeks between shots and were invited later. However, once appropriate
144 considerations have been made regarding vaccine and case data, it is possible to incorporate
145 (time-dependent) vaccination coverage rates in endemic-epidemic modelling. To appropriately
146 account for the remaining (unvaccinated) number of susceptibles, use of the log proportion of
147 unvaccinated cases is recommended following Herzog et al. (2011). This is also the approach
148 Kucharski and colleagues have utilised in their endemic-epidemic model for measles which in-
149 cluded vaccination (Robert et al., 2021).

150 **Interpretability (Kucharski)** It is true that there is a balance between what data allows us
151 to fit and how realistic and interpretable our model is. The benefit of the endemic-epidemic
152 modelling framework is that it allows us to examine the spread of disease across age groups with
153 flexible statistical techniques. The first instance of such a multivariate model is Knorr-Held and
154 Richardson (2003) which investigated the spatio-temporal dynamics of meningococcal disease.
155 Compartmental models are easier to interpret, but more difficult to apply to surveillance data (see
156 Held et al., 2006; Paul et al., 2008, for further discussion).

REFERENCES

- 158 Aaby, P., Thoma, H., and Dietz, K. (2021). Measles in the European past: Outbreak of severe
159 measles in an isolated German village, 1861. *SSRN*.
- 160 Alleman, T. W., Vergeynst, J., Torfs, E., Illana G., D., Nopens, I., and Baetens, J. M. (2020).
161 A deterministic, age-stratified, extended SEIRD model for assessing the effect of non-
162 pharmaceutical interventions on SARS-CoV-2 spread in Belgium [pre-print]. *medRxiv*.
- 163 Arregui, S., Aleta, A., Sanz, J., and Moreno, Y. (2018). Projecting social contact matrices to
164 different demographic structures. *PLOS Computational Biology*, 14(12):1–18.
- 165 Bauer, C., Wakefield, J., Rue, H., Self, S., Feng, Z., and Wang, Y. (2016). Bayesian penal-
166 ized spline models for the analysis of spatio-temporal count data. *Statistics in Medicine*,
167 35(11):1848–1865.
- 168 Bracher, J. and Held, L. (2020a). Endemic-epidemic models with discrete-time serial interval
169 distributions for infectious disease prediction. *International Journal of Forecasting*.
- 170 Bracher, J. and Held, L. (2020b). A marginal moment matching approach for fitting endemic-
171 epidemic models to underreported disease surveillance counts. *Biometrics*, pages 1–13.
- 172 Brooks-Pollock, E., Danon, L., Jombart, T., and Pellis, L. (2021). Modelling that shaped the early
173 COVID-19 pandemic response in the UK. *Philosophical Transactions of the Royal Society B:
174 Biological Sciences*, 376(1829):20210001.
- 175 Chambers, C. (2019a). The registered reports revolution – lessons in cultural reform. *Signifi-
176 cance*, 16(4):23–27.
- 177 Chambers, C. (2019b). What’s next for registered reports. *Nature*, 573(7773):187–189.
- 178 Feehan, D. and Mahmud, A. (2020). Quantifying interpersonal contact in the United States
179 during the spread of COVID-19: first results from the Berkeley Interpersonal Contact Study.
180 *medRxiv*.

- 181 Grimée, M., Bekker-Nielsen Dunbar, M., Hofmann, F., and Held, L. (2021). Modelling the effect
182 of a border closure between Switzerland and Italy on the spatiotemporal spread of COVID-19
183 in Switzerland. *medRxiv*.
- 184 Groseclose, S. L. and Buckeridge, D. L. (2017). Public health surveillance systems: Recent
185 advances in their use and evaluation. *Annual Review of Public Health*, 38:57–79.
- 186 Held, L., Hens, N., O’Neill, P., and Wallinga, J., editors (2020). *Handbook of Infectious Disease*
187 *Data Analysis*. Chapman & Hall/CRC Handbooks of Modern Statistical Methods.
- 188 Held, L., Hofmann, M., Höhle, M., and Schmid, V. (2006). A two-component model for counts
189 of infectious diseases. *Biostatistics*, 7(3):422–437.
- 190 Held, L. and Meyer, S. (2020). Forecasting based on surveillance data. In Held, L., Hens, N.,
191 O’Neill, P., and Wallinga, J., editors, *Handbook of infectious disease data analysis*, chapter 25,
192 pages 509–528. Chapman & Hall/CRC Handbooks of Modern Statistical Methods.
- 193 Held, L., Meyer, S., and Bracher, J. (2017). Probabilistic forecasting in infectious disease epi-
194 demiology: the 13th Armitage lecture. *Statistics in Medicine*, 36(22):3443–3460.
- 195 Herzog, S. A., Paul, M., and Held, L. (2011). Heterogeneity in vaccination coverage explains
196 the size and occurrence of measles epidemics in German surveillance data. *Epidemiology &*
197 *Infection*, 139(4):505–515.
- 198 Höhle, M. and an der Heiden, M. (2014). Bayesian nowcasting during the STEC O104:H4
199 outbreak in Germany, 2011. *Biometrics*, 70(4):993–1002.
- 200 Hulth, A., N., A., Ethelberg, S., Dreesman, J., Faensen, D., van Pelt, W., and Schnitzler, J.
201 (2010). Practical usage of computer-supported outbreak detection in five European countries.
202 *Eurosurveillance*, 15(36):19658.
- 203 Jarvis, C. I., Gimma, A., van Zandvoort, K., Wong, K. L. M., CMMID COVID-19 work-
204 ing group, and Edmunds, W. J. (2020a). The impact of local and national restrictions in

205 response to COVID-19 on social contacts in the UK: a longitudinal natural experiment.
206 <https://cmmid.github.io/topics/covid19/comix-national-local-restrictions-in-the-UK.html>.

207 Jarvis, C. I., Van Zandvoort, K., Gimma, A., Prem, K., CMMID COVID-19 working group,
208 Klepac, P., Rubin, G. J., and Edmunds, W. J. (2020b). Quantifying the impact of physical
209 distance measures on the transmission of COVID-19 in the UK.

210 Knorr-Held, L. and Richardson, S. (2003). A hierarchical model for space-time surveillance
211 data on meningococcal disease incidence. *Journal of the Royal Statistical Society: Series C*
212 *(Applied Statistics)*, 52(2):169–183.

213 Latsuzbaia, A., Herold, M., Bertemes, J.-P., and Mossong, J. (2020). Evolving social contact
214 patterns during the COVID-19 crisis in Luxembourg. *PLOS ONE*, 15(8):1–13.

215 Mistry, D., Litvinova, M., Pastore y Piontti, A., Chinazzi, M., Fumanelli, L., Gomes, M. F. C.,
216 Haque, S. A., Liu, Q.-H., Mu, K., Xiong, X., Halloran, M. E., Longini Jr., I. M., Merler, S.,
217 Ajelli, M., and Vespignani, A. (2020). Inferring high-resolution human mixing patterns for
218 disease modeling. *arXiv*. <https://arxiv.org/abs/2003.01214>.

219 Paul, M., Held, L., and Toschke, A. M. (2008). Multivariate modelling of infectious disease
220 surveillance data. *Statistics in Medicine*, 27(29):6250–6267.

221 Polonsky, J. A., Baidjoe, A., Kamvar, Z. N., Cori, A., Durski, K., Edmunds, W. J., Eggo, R. M.,
222 Funk, S., Kaiser, L., Keating, P., de Waroux, O. I., Marks, M., Moraga, P., Morgan, O., Nou-
223 vellet, P., Ratnayake, R., Roberts, C. H., Whitworth, J., and Jombart, T. (2019). Outbreak
224 analytics: a developing data science for informing the response to emerging pathogens. *Philo-
225 sophical Transactions of the Royal Society B: Biological Sciences*, 374(1776):20180276.

226 Prem, K., Liu, Y., Russell, T., Kucharski, A. J., Eggo, R. M., Davies, N., Centre for the Mathe-
227 matical Modelling of Infectious Diseases COVID-19 Working Group, Jit, M., and Klepac, P.
228 (2020). The effect of control strategies that reduce social mixing on outcomes of the COVID-
229 19 epidemic in Wuhan, China. *Lancet Public Health*.

- 230 Ray, E. L., Sakrejda, K., Lauer, S. A., Johansson, M. A., and Reich, N. G. (2017). Infectious
231 disease prediction with kernel conditional density estimation. *Statistics in Medicine*, 36:4908–
232 4929.
- 233 Robert, A., Kucharski, A. J., and Funk, S. (2021). The impact of local vaccine coverage and
234 recent incidence on measles transmission in France between 2009 and 2018. *medRxiv*.
- 235 Schwab, S. and Held, L. (2020). Science after Covid-19: Faster, better, stronger? *Significance*,
236 17(4):8–9.
- 237 Stojanović, O., Leugering, J., Pipa, G., Ghozzi, S., and Ullrich, A. (2019). A bayesian monte
238 carlo approach for predicting the spread of infectious diseases. *PLOS ONE*, 14(12).
- 239 Van den Akker, O., Weston, S. J., Campbell, L., Chopik, W. J., Damian, R. I., Davis-Kean, P.,
240 Hall, A. N., Kosie, J. E., Kruse, E. T., Olsen, J., Richie, S., Valentine, K., van 't Veer, A.,
241 and Bakker, M. (2020). Preregistration of secondary data analysis: A template and tutorial.
242 *PsyArXiv*.
- 243 van Leeuwen, E., PHE Joint modelling group, and Sandmann, F. (2020). Augmenting contact
244 matrices with time-use data for fine-grained intervention modelling of disease dynamics: A
245 modelling analysis [pre-print]. *medRxiv*.
- 246 Willem, L., Hoang, T. V., Funk, S., Coletti, P., Beutels, P., and Hens, N. (2020). SOCRATES:
247 An online tool leveraging a social contact data sharing initiative to assess mitigation strategies
248 for COVID-19. *BMC Research Notes*, 13(1):293.