

1 **Evaluating the use of social contact data to produce age-specific** 2 **forecasts of SARS-CoV-2 incidence**

3 James D Munday^{*1,2}, Sam Abbott^{1,2}, Sophie Meakin^{1,2} & Sebastian Funk^{1,2}

4 1. *Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene*
5 *and Tropical Medicine, UK*

6 2. *Department of Infectious Disease Epidemiology, London School of Hygiene and*
7 *Tropical Medicine, UK*

8 [*james.munday@lshtm.ac.uk](mailto:james.munday@lshtm.ac.uk)

9 **Abstract**

10 Short-term forecasts can provide predictions of how an epidemic will change in the near
11 future and form a central part of outbreak mitigation and control. Renewal-equation based
12 models are increasingly popular. They infer key epidemiological parameters from historical
13 epidemiological data and forecast future epidemic dynamics without requiring complex
14 mechanistic assumptions. However, these models typically ignore interaction between age-
15 groups, partly due to challenges in parameterising a time varying interaction matrix. Social
16 contact data collected regularly by the CoMix survey during the COVID-19 epidemic in
17 England, provide a means to inform interaction between age-groups in real-time.

18 We developed an age-specific forecasting framework and applied it to two age-stratified
19 time-series: incidence of SARS-CoV-2 infection, estimated from a national infection and
20 antibody prevalence survey; and, reported cases according to the UK national COVID-19
21 dashboard. Jointly fitting our model to social contact data from the CoMix study, we inferred
22 a time-varying next generation matrix which we used to project infections and cases in the
23 four weeks following each of 29 forecast dates between October 2021 and November 2022.
24 We evaluated the forecasts using proper scoring rules and compared performance with
25 three other models with alternative data and specifications alongside two naive baseline
26 models.

27 Overall, incorporating age-interaction improved forecasts of infections and the CoMix-data-
28 informed model was the best performing model at time horizons between two and four
29 weeks. However, this was not true when forecasting cases. We found that age-group-
30 interaction was most important for predicting cases in children and older adults. The contact-
31 data-informed models performed best during the winter months of 2020 - 2021, but
32 performed comparatively poorly in other periods. We highlight challenges regarding the
33 incorporation of contact data in forecasting and offer proposals as to how to extend and
34 adapt our approach, which may lead to more successful forecasts in future.

35 Introduction

36 Effective epidemic response relies on accurate infection surveillance to provide status
37 updates which support decision makers[1]. Surveillance data can be enhanced by estimating
38 key epidemiological parameters in real-time such as the growth rate and time-varying
39 reproduction number (R_t) and by generating short-term-forecasts of incidence of infection,
40 hospitalisation and mortality[2–4]. These provide estimates of the current and future
41 epidemic trajectory to public health decision makers. As such a host of approaches have
42 been developed to make short term epidemiological forecasts. A popular genre of
43 methodology for infectious disease forecasts are renewal-equation based ‘semi-mechanistic’
44 models [2,4–6], which infer key epidemiological parameters from historical time-series data,
45 in particular changes in transmission intensity, and use them to forecast future epidemic
46 dynamics without requiring the more detailed assumptions and complex mathematical
47 framework involved in ‘fully-mechanistic’ models (e.g. compartmental or agent based
48 models).

49 Age has been shown to be an important factor in both transmission risk [7,8] and severity of
50 disease [9–11] caused by SARS-CoV-2. This is not unique to the COVID epidemic. In the
51 past, epidemiological analysis and modelling have shown variability and homophily in
52 transmission by age to have important implications for the dynamics of infection[12–15].
53 Moreover, age distribution of infection has important implications for the potential burden of
54 disease as infection moves between age groups, who are more and less prone to severe
55 illness and death[7,8,16].

56 Although age-specific forecasts are desirable to better understand the risk to particularly
57 vulnerable groups, due to variance in prevalence of infection between age-groups, accurate
58 estimates of future incidence might require risk of transmission between age groups to be
59 captured effectively. However, the high dimensionality of the problem means that estimating
60 an age-interaction matrix is not possible from epidemiological data alone [17]. Instead, much

61 infectious disease dynamics research in the past 30 years has made assumptions in line
62 with the social contact hypothesis [17]. It states that the rate of transmission of directly
63 infectious agents is proportional to the population-level rate of social contact between
64 population groups. This hypothesis is the basis for age-structured mixing assumptions in
65 many mathematical models. Such models are generally parameterised from data gathered in
66 social contact surveys[15,18], which typically ask participants to report their social contacts
67 from a fixed period in the recent past, e.g. the last 24 hours. Participants are also asked
68 about the characteristics of their contacts at each contact event, usually including age[15]. A
69 key challenge to the use of historically collected contact data has been of their temporal and
70 geographical generalisability. This is especially true when non-pharmaceutical interventions
71 (NPIs) are in effect, potentially drastically changing the contact behaviour of the general
72 public. The variability in behaviour with time and age during a pandemic makes
73 parameterisation of age-specific-real-time-models particularly challenging as up-to-date
74 information on interaction is essential for time-varying parameterisation of the model.

75 During the COVID-19 pandemic, as a means to monitor the behaviour of the general public
76 relevant to transmission and provide insight into risk posed to vulnerable populations, a
77 number of studies were conducted to survey social contacts at a frequency and scale not
78 seen previously. One example is the CoMix study, which collected contact data weekly
79 between March 2020 to March 2022 in 19 European countries[19–22]. The UK arm of the
80 study, which involved a survey of greater than 5000 participants, was the first to launch and
81 most complete in terms of data collected over this period. This regularly collected contact
82 data provides a means to parameterise models with temporally and geographically relevant
83 estimates of social interaction, and an opportunity to evaluate how incorporating such data
84 into a real-time analysis framework impacts forecast performance at different scales.

85 Existing studies of forecasting performance[5,6,23] have focused on age-agnostic numbers
86 of cases, hospitalisations and deaths. Probabilistic forecasts can be robustly assessed using
87 proper scoring rules [24]. Although these methods have been popular for some time in other

88 fields, they have only recently been applied to epidemic forecasts[5,6,23]. To the authors
89 knowledge one such evaluation has previously been made [25] of age-stratified epidemic
90 forecasts however, the study by Held et. al. used historical contact data to parameterize
91 interaction between age-groups and evaluated at a population level by summing age-specific
92 forecasts. To our knowledge there has been no evaluation of the use of the regularly
93 collected age-stratified contact data in comparison to other approaches to make short term
94 forecasts at an age-group specific level.

95 Here we present age-specific forecasts in the UK, with the aim of understanding whether
96 incorporating the weekly collected social contact data improves the predictive ability
97 compared to ignoring this interaction. We incorporated data from the CoMix study in a semi-
98 mechanistic forecasting framework and applied this to case numbers, as the most commonly
99 tracked metric for COVID-19 dynamics in the UK throughout the pandemic. We further
100 applied it to infection incidence estimated from a weekly cross-sectional household survey of
101 infection [26,27] in order to better understand the influence of reporting patterns on results.
102 To quantify the relative benefits of incorporating interaction between age groups and specific
103 contact data into forecasts we compared three models with interaction between age groups
104 with an equivalent model with no such interaction and evaluated the models against two
105 naive baseline models.

106 **Materials and Methods**

107 Study overview

108 To establish the relative benefit of incorporating interaction between age-groups in short-
109 term epidemiological forecasts, we implemented four age-stratified semi-mechanistic
110 models, which each estimate a time-varying Next Generation Matrix (NGM). This matrix is
111 inferred as the interaction matrix between age-groups under the assumption that all
112 infections in each age group are informed by the sum of past infections in all age-groups
113 weighted by the distribution of time between infections - the generation interval distribution

114 and the NGM. Two of the models included interaction between age groups, one of which
115 was informed by contact data from the CoMix study (regularly collected during the period of
116 study). We also evaluate the same model using data collected in the POLYMOD survey
117 (single survey performed in 2008). In the fourth model the interaction was estimated entirely
118 from historical epidemiological data. We compared these models with a fourth model which
119 allowed no interaction between age-groups.

120 We applied this to reported cases, as a commonly available quantity for forecasting epidemic
121 dynamics[5]. This, however, incurs a secondary challenge due to potential variability in
122 reporting of cases by age and over the course of an epidemic, which may serve to
123 complicate our interpretation of the application of contact data to forecasts. Hence, to isolate
124 the impact of incorporating contact data we chose to additionally apply the models to
125 estimated infection incidence from a repeated cross-sectional household survey of
126 infections.

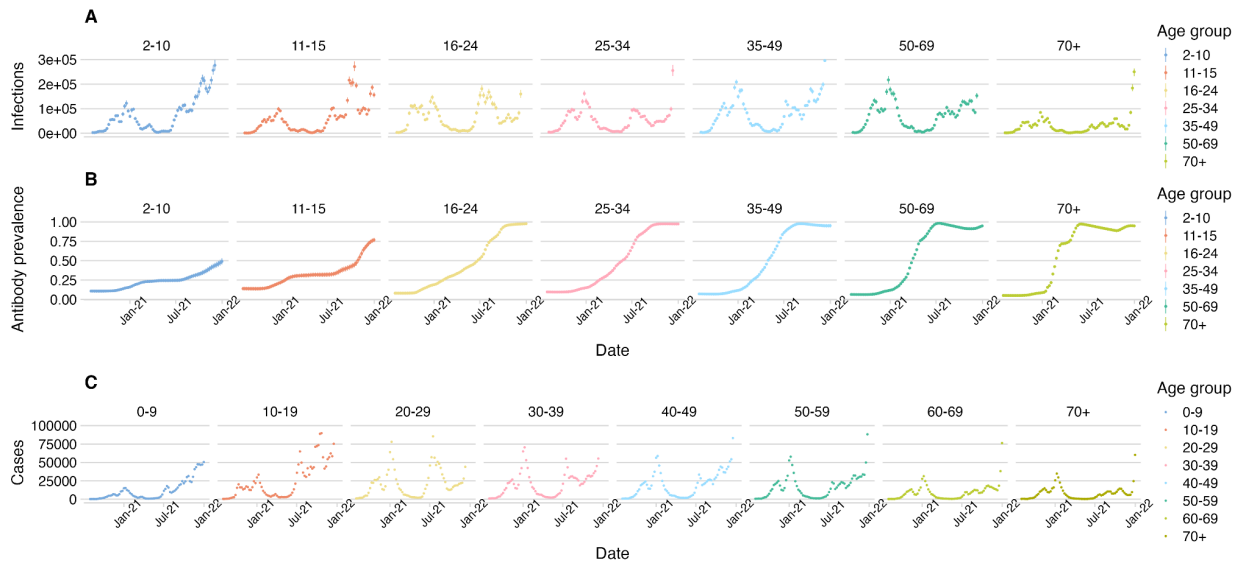
127 We forecast weekly reported cases using the data from the UKHSA Covid-19 dashboard.
128 For convenience we used the full case time-series aggregated to weekly incidence and
129 truncated at different forecast dates, rather than the data available on each forecast date.
130 Although this does not give a full picture of the real-time applicability and performance of the
131 model, it avoids complications in delays in gathering case reports which require additional
132 treatment prior to application of a forecasting model such as truncation of the most recent
133 data or now-casting[28]. Secondly, we applied the models to estimates of weekly incidence
134 of infection estimated[26] from national infection prevalence data, again with the full final
135 data set truncated at each forecast date rather than snapshots available at the time. To
136 further isolate the role of contact data in the forecasts of infections, we used weekly age-
137 stratified estimates of antibody prevalence to inform age-specific susceptibility.

138 Data

139 We accessed daily, age-stratified, case data from the UK COVID-19 dashboard [29] on 11th
140 May 2022. We aggregated this data to weekly incidence by taking the 7 day sum from the
141 previous 7 days, aligned such that the weekly data is reported on the proposed forecast
142 dates, to forecast weekly case counts in future weeks. The case reports were grouped in
143 seven decade groups between zero and 69 years with a single group for over 70 year olds
144 (0-9, 10-19, ..., 60-69, 70+).

145 We accessed aggregates of SARS-CoV-2 infection prevalence and antibody prevalence
146 collected as part of the COVID-19 infection survey(CIS) through the CIS Website[27] on 18th
147 March 2022. We used data covering the period between August 2020 and January 2022 to
148 estimate weekly infection incidence and antibody prevalence for seven age-groups (2-10,
149 11-15, 16-24, 25-34, 35-49, 49-69 and 70+). In addition to the CIS data, we used vaccination
150 data published by the National Health Service and accessed via the UK coronavirus
151 dashboard[30] on the same date.

152 We generated SARS-CoV-2 infection incidence and antibody prevalence time-series for the
153 period between August 2020 and January 2022 using an approach described
154 elsewhere[26,31](Figure 1). To establish a weekly time-series of infections we took the sum
155 of incident infections in each week on a sample-by-sample basis and calculated the credible
156 intervals from the resultant sum. To establish a weekly time series of antibodies we took the
157 antibody prevalence on the last day of each week and calculated credible intervals from the
158 full posterior sample.



159

160 **Fig. 1. Estimated incidence of A) infection, B) antibody prevalence and C) case reports by age-**
 161 **group.**

162 We combined these data with social contact data collected as part of the CoMix social
 163 contact survey (CoMix)[19,32], a multinational, weekly, cross-sectional survey of social
 164 contacts. We used published weekly contact matrices from the UK arm of CoMix, generated
 165 under the framework described previously[20].

166 Transmission Model

167 We extended the concept of the Next Generation Matrix to include transmission interval
 168 distributions (the generation interval for infections, and the interval between a positive test in
 169 infector and infectee for cases). Here, the number of incident cases or infections $I(t)$ at time t
 170 was given by the sum of the products of the next generation matrix \mathbb{N} and the age-stratified
 171 vector of cases or infections on dates between $t - s_{max}$ and $t - 1$, weighted by the
 172 transmission interval distribution $w(s)$.

$$I(t) = \sum_{s=1}^{s_{max}} w(s) \times \mathbb{N}(t - s)I(t - s) \quad (1)$$

173

174 where S_{max} is a fixed upper-limit of the transmission interval distribution, set to 4 weeks and
175 $w(s)$ is assumed to follow a discretised log-normal distribution with time since the primary
176 event (infection or positive test of the infector):

$$177 \quad w(s) = (F_{LNorm}(s, w_\mu, w_\sigma) - F_{LNorm}(s-1, w_\mu, w_\sigma)) / F_{LNorm}(S_{max}, w_\mu, w_\sigma) \quad (2)$$

178 where F_{LNorm} is the cumulative distribution function of the log-normal distribution with
179 parameters w_μ and w_σ . Under the social contact hypothesis[17], the next generation matrix
180 is calculated by multiplying the contact matrix, $C(t)$ quantifying the mean number of contacts
181 between age groups, with vectors of age-specific susceptibility (s) and infectiousness (i),
182 where each element, s_a and i_a give the specific susceptibility and infectiousness of age group
183 a [13].

$$184 \quad N(t) = \text{diag}(s)C(t)\text{diag}(i) \quad (3)$$

185 We assumed that age specific infectiousness, i , is inherent and unrelated to time varying
186 factors associated with the epidemic. We assumed that age-specific susceptibility included
187 two components:

$$188 \quad s_a = s_{ab,a} s_{inh,a} \quad (4)$$

189 The first (s_{ab}) is drawn from age-specific immunity to infection, which is informed by age-
190 specific antibody prevalence. We used a leaky definition of antibody effectiveness in line with
191 the definition used in the estimation of the infection and antibody timeseries:

$$192 \quad s_{ab,a} = 1 + (\Phi - 1)A_a(t) \quad (5)$$

193 Where Φ is the effectiveness of antibodies in preventing infections in an exposed member of
194 the population. The second component (s_{inh}) is due to an age-correlated variation in inherent
195 susceptibility to infection and unrelated to time-varying factors associated with the epidemic.
196 Both s_{inh} and i were assumed to remain constant in time, such that all of the variation in the
197 next generation matrix by time is governed by changes in contacts and estimated antibody

198 derived immunity. Both S_{inh} and i were fit as random effects in a hierarchical framework
199 (Table 1).

200 Parameter estimation and forecasting

201 To allow variation in parameter values over the course of the study period, we parameterised
202 the model with the estimated antibody prevalence and contact matrices and fitted it to 8
203 weeks of weekly estimated infection or reported case data prior to the forecast date. We
204 fitted the model using Hamiltonian Monte Carlo, implemented in the Stan probabilistic
205 programming language[33] assessing the convergence of each model by monitoring
206 transition divergence and chain mixing.

207 We fitted to the mean infection time series $I_\mu(t)$ under the likelihood:

$$208 \quad I_\mu(t) \sim \text{normal}\left(\sum_{s=1}^{s_{max}} w(s)N(t-s)I_\mu(t-s), \sigma'_I(t)\right) \quad (6)$$

209 Where σ'_I is the overall uncertainty in the modelled infections, combining the time-varying
210 inherent uncertainty in the NGM model, σ_I , and the standard deviation of the infection
211 estimates, I_σ ,

$$212 \quad \sigma'_I(t) = \sqrt{I_\sigma^2 + \sigma_I(t)^2} \quad (7)$$

213 σ_I is constructed at each time point from the estimated coefficient of variation CV_I and
214 infection incidence such that:

$$215 \quad \sigma_I(t) = CV_I I_\mu(t) \quad (8)$$

216 which ensures the uncertainty scales with the magnitude of the infection incidence
217 estimates. We fit to the case time series $c(t)$ under the likelihood:

$$218 \quad c(t) \sim \text{normal}\left(\sum_{s=1}^{s_{max}} w(s) \times N(t-s)c(t-s), \sigma_c(t)\right) \quad (9)$$

219 Where σ_c is the modelled uncertainty in cases and is constructed from the estimated
220 coefficient of variation CV_c at each time point such that:

$$221 \quad \sigma_I(t) = CV_c c_\mu(t) \quad (10)$$

222 which ensures the uncertainty scales with the magnitude of the reported case incidence. To
223 incorporate the contact data in the CoMix based model we jointly fit the contact matrices
224 under the likelihood:

$$225 \quad C_{ab}(t) \sim \text{normal}(C_{ab}, \sigma'_{cm,ab}) \quad (11)$$

$$226 \quad \sigma'_{cm,ab} = \sqrt{C_{\sigma,ab}^2 + \sigma_{cm}^2} \quad (12)$$

227 and $C_{\sigma,ab}$ is the estimated standard deviation of the measured contact rate and σ_{cm} is the
228 uncertainty in the fitted contacts.

229 Each of the models estimated a NGM which varied over the 8 weeks of prior data only by
230 changes in the estimated contact matrices and antibody inferred immunity, whilst the
231 inherent susceptibility and infectiousness vectors were assumed constant for the whole
232 modelled period. However, as each forecast date was modelled independently, all
233 parameters were able to vary between forecasts.

234 We used uninformative priors (Table 1) for the contact rate between each pair of age groups
235 (C_{ab}), model uncertainty parameters (CV_I, CV_c, σ_{cm}) and antibody protection (Φ). Antibody
236 prevalence priors were set to the distribution of the estimate provided by the model used to
237 estimate incidence (*inc2prev*)[31] and relative susceptibility and infectiousness vector
238 elements were set such that the Secondary Attack Rate (SAR) was roughly half that of
239 estimates of Household SAR in literature[34]—which aimed to account for reduced risk of
240 transmission to known contacts outside the household. The prior for the log-mean (w_{mu}) and
241 log-standard-deviation (w_{sigma}) of the transmission interval had a mean and standard
242 deviation of 5 days to reflect the broad distribution of transmission intervals recorded in

243 literature [35–37], these were converted to the appropriate log-parameters for the log-normal
244 framework in equation 2, and their prior was set to be normally distributed with a standard
245 deviation of 20% of the mean.

246 We used posterior distributions of the parameters (Table 1) to project infections and cases
247 forwards up to four weeks after each forecast date. We note that contact data directly
248 relevant to the dates forecasted would not be known on the forecast date, so we used the
249 contact data corresponding to the week of the forecast date itself, assuming that these also
250 reflected contacts in the following week. For the case forecasts we offset the contact data by
251 7 days to account for delay between infection and specimen date and used the generation
252 interval as a proxy of the test-to-test distribution[38], which is consistent with a 5 day
253 incubation period and a 2 day report delay[39].

254 **Table 1.** Model parameters and priors

parameter	Symbol	Prior
Antibodies	$A_a(t)$	$A(t) \sim \text{normal}(A_{a,\mu}(t), A_{a,\sigma}(t))T[0, 1]$
Antibody protection	Φ	$\phi \sim \text{gamma}(2, 2)T[0, 1]$
Generation interval distribution log-mean and log-variance	$w(t)$	$w_{mean} = 5/7$ $w_{sd} = 5/7$ $w_{\sigma,prior} = \log(((w_{sd}^2)/(w_{mean}^2)) + 1)$ $w_{\mu,prior} = \log(w_{mean}) - (w_{\sigma,prior})/2$ $w_{\mu} \sim \text{normal}(w_{\mu,prior}, w_{\mu,prior}/5)T[0,]$ $w_{\sigma} \sim \text{normal}(w_{\sigma,prior}, w_{\sigma,prior}/5)T[0,]$
inherent susceptibility	s_{inh}	$s_{\mu}^h \sim \text{Beta}(24, 24)$ $s_{sd}^h \sim \text{normal}(0.1, 0.02)T[0,]$ $s'_{inh,a} \sim \text{normal}(0, 1)$ $s_{inh,a} = s_{\mu}^h + s_{sd}^h s'_{inh,a}$
inherent infectiousness	i_a	$i_{\mu}^h \sim \text{Beta}(12, 4)$ $i_{sd}^h \sim \text{normal}(0.1, 0.02)T[0,]$ $i'_a \sim \text{normal}(0, 1)$ $i_a = i_{\mu}^h + i_{sd}^h i'_a$
Contact matrices	C	$C_{ab} \sim \text{gamma}(2, 2)$
Uncertainty in infections	CV_I	$CV_I \sim \text{normal}(0.05, 0.025)T[0,]$
Uncertainty in cases	CV_c	$CV_c \sim \text{normal}(0.05, 0.01)T[0,]$
Uncertainty in contacts	σ_{cm}	$\sigma_{cm} \sim \text{normal}(0.05, 0.025)T[0,]$
Where $T[a, b]$ indicates distribution is truncated between the values a and b		

255

256 Model evaluation

257 We evaluated the performance of the NGM models (CoMix-data, No-contact-data and No-
258 interaction) across 29 forecast dates between October 2020 and December 2021. We chose
259 this period as there was major disruption to the CoMix survey during July 2020 and following
260 changes in the survey in June 2020. We excluded dates after December 2021 due to the
261 complication of the emergence of the Omicron variant, which has been shown to evade
262 immunity to a greater extent than earlier variants[40], complicating our interpretation of
263 antibody prevalence as a mix of omicron-specific and previously acquired antibodies persist
264 in the population.

265 We evaluated the forecasts against the reported number of cases or mean estimated
266 number of infections in the week forecasted, for case and infection forecasts respectively.
267 We evaluated the forecasts based on Continuous Ranked Probability Score (CRPS) and a
268 measure of bias (see appendix for definitions) each implemented using the *scoringutils* R
269 package [41]. The CRPS measures the ‘distance’ of the predictive distribution to the
270 observed data-generating distribution, hence a lower score indicates more accurate
271 predictions and therefore a higher performing model. The bias measures the tendency for a
272 model to over (positive value) or under (negative value) predict the incidence in its
273 projections, hence a bias of zero is optimal.

274 We also assessed the models calibration by evaluating the central interval coverage
275 (coverage) of each model's forecast (Proportion of incidence points which fell in the ranges
276 projected by the forecast model's posterior distribution of future cases).

277 To provide a comparator as a lower bound of performance, we also evaluated two baseline
278 models. These baselines were intended to represent naive assumptions, which may be
279 applied without the use of a model. The first baseline assumed no change in incidence from
280 the day the forecast was made. The second calculated the change in incidence between the
281 forecast date and each week within the four week forecast horizon, the rate of change is

282 projected as an exponential extrapolation based on the previous two weeks of data. Both
283 baselines were modelled without uncertainty and, consequently, the CRPS reduced to the
284 mean absolute error. To provide a clear comparison of performance with and without
285 interaction between age-groups, we provide all CRPS scores relative to the score of the no-
286 interaction model (rCRPS).

287 As well as the overall performance of each forecasting model, we also evaluated the
288 forecasts by grouping forecasts made by each model in two ways. Firstly, we aggregated the
289 forecasts by age-group—showing the relative performance of the models when forecasting
290 incidence in particular age categories. Secondly, to evaluate how performance changed over
291 the course of the analysis period, we scored the forecasts separately for seven key phases
292 of the pandemic (Table 2). For this we used the phases used in Gimma, et. al. [19] which
293 overlapped with our analysis, with the addition of two phases that were not covered by the
294 previous CoMix work. Due to the small number of weeks covered by ‘Christmas’ and
295 ‘Lockdown 3 schools open’ we combined these with ‘Lockdown 3’ and ‘Lockdown 3 Easing’
296 respectively.

297 **Table 2.** Pandemic period names and dates

Period	Start date	End date
Lockdown 2	2020-11-05	2020-12-02
Lockdown 2 Easing	2020-12-03	2020-12-19
Christmas	2020-12-20	2021-01-04
Lockdown 3	2021-05-01	2021-03-08
Lockdown 3 Schools open	2021-03-09	2021-03-28
Lockdown 3 easing	2021-03-29	2021-09-30
Opening up	2021-10-01	2021-11-24

298

299

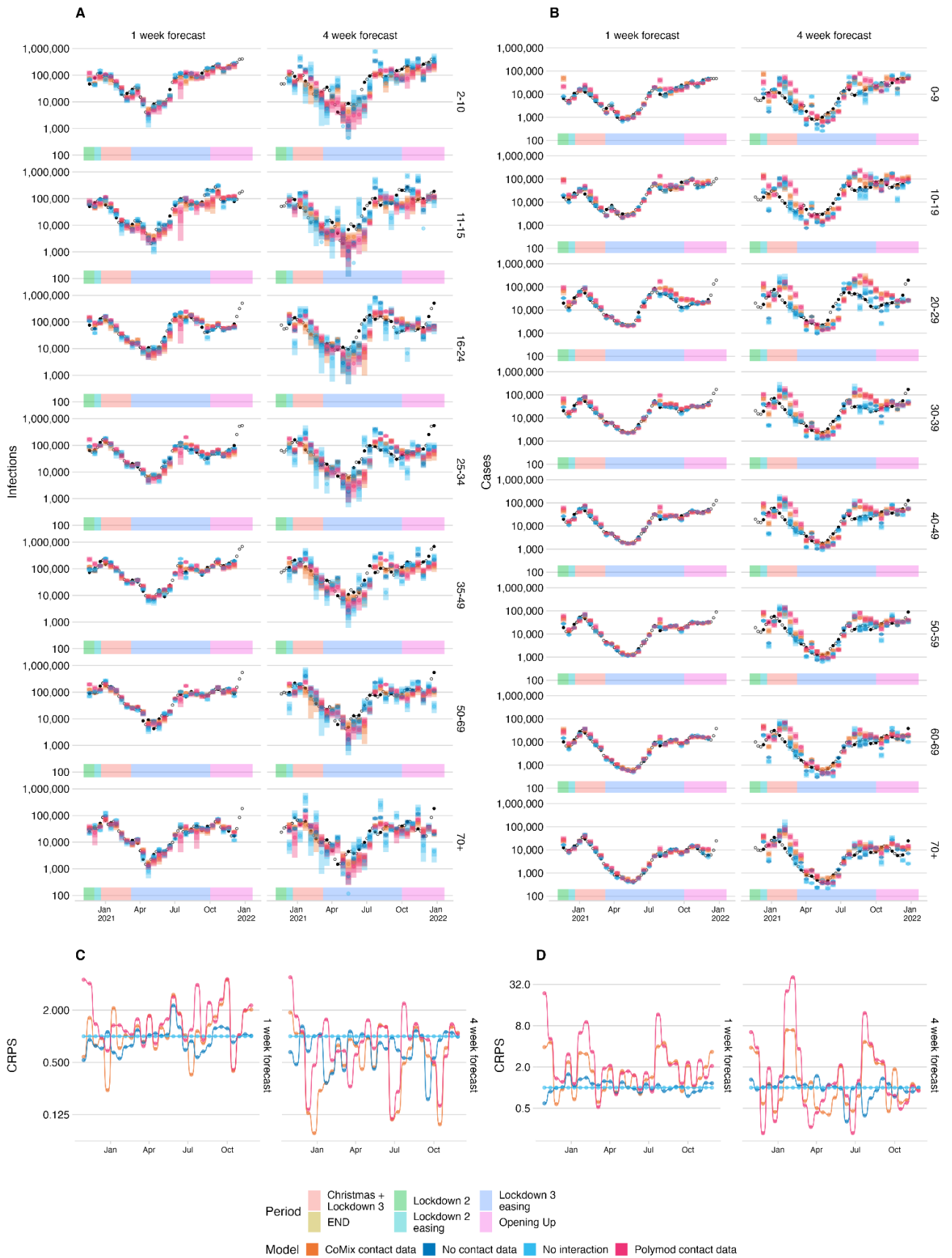
300 Age-specific transmission parameters

301 Finally, to compare the implicit assumptions within the models we applied, we assessed how
302 the values of the relative susceptibility and infectiousness parameters s and i varied over the
303 pandemic. To provide an interpretable quantification of these parameters, we used the age-
304 specific values estimated in the model to calculate ratios of susceptibility and infectiousness
305 of younger adults, and older adults relative to that of children. Due to the different age-
306 stratification in the data available, the broader age-bands here varied between case forecast
307 models and infection forecast models: Children were defined as up to 15 for infections and
308 up to 19 for cases, younger adults were defined as 16-49 for infections and 20-49 for cases
309 and older adults were defined as over 50 in both instances.

310 **Results**

311 Forecasts

312 We made forecasts with a horizon of one, two, three and four weeks at fortnightly intervals
313 (Figure 2, Figures S1 - S6) between 30th October 2020 and 26th November 2021 (29
314 forecast dates). Visually the forecasts deviate more from the true data at longer forecast
315 horizons.



317 **Fig. 2** A) and B) Infections and cases, respectively, forecast using the CoMix-data based next
318 generation model, the no-contact-data, no-interaction and POLYMOD-data data based model, for
319 each age group (top to bottom) and forecast horizon (left to right). projected infections from each
320 model (coloured bars) and black points show infection estimates and reported cases in plots A and B
321 respectively. The estimates being forecast on each axis are shown as solid points; those not being
322 forecast are shown as rings. C) and D) show the continuous ranked probability score relative to the
323 score of the “no interaction” model for each forecast date, calculated from the Infection and Case
324 forecasts respectively.

325 Model Evaluation

326 To assess the relative performance of each of the models for different forecast horizons, we
327 calculated evaluation metrics separately for each forecast horizon across all forecast dates
328 (Figure 3 A, Table S1). For an alternative approach using multivariate evaluation across age
329 groups and time horizons, see Held et al. (2017) [25].

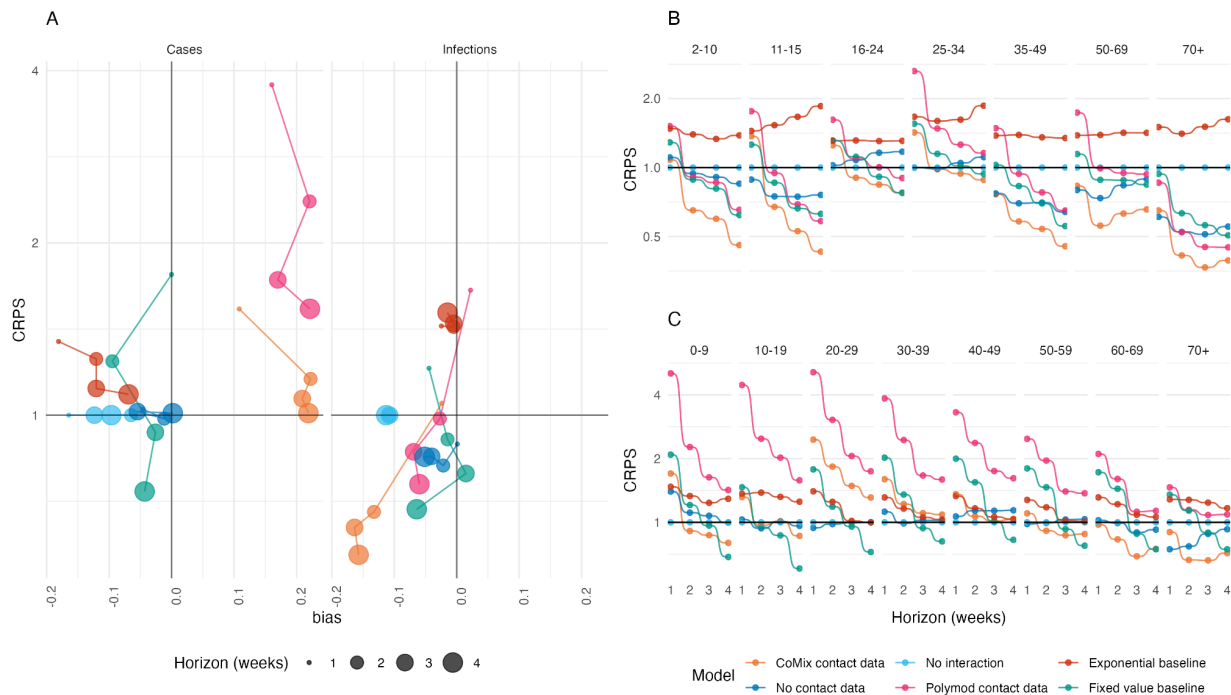
330 When forecasting case reports from the UK Covid-19 dashboard [29], the non-interaction model
331 performed better than any of the models which allowed interaction. The next best model was the
332 model with no contact data which performed similarly to the no-interaction model, particularly at
333 longer time horizons with rCRPS of 1.02, 0.99, 1.01 and 1.01 (relative to the no-interaction
334 model) respectively at horizons of one to four weeks in ascending order. The two models that
335 incorporated contact data both performed poorly when considering the CRPS relative to the no-
336 interaction model, with the Polymod model performing the worst. However, the relative
337 performance of both models improved at longer time horizons, with the CoMix model performing
338 similarly to the other NGM models at four week horizons (rCRPS of 1.53 and 3.78 at one week
339 horizon reducing to 1.01 and 1.53 at four week horizon for the CoMix and Polymod data models
340 respectively). Both CoMix and Polymod forecasts had a substantial positive bias, showing that,
341 on average, they over-predicted cases with bias between 0.15 and 0.25. The other models tended
342 to under predict cases by a smaller margin (between 0 and 0.18). The exponential baseline
343 performed worse than the no-interaction model at all forecast horizons (rCRPS 1.35, 1.25, 1.11

344 and 1.09 at one to four weeks in ascending order). The fixed value baseline initially performed
345 second to worst for one week forecasts (rCRPS=1.76) but improved as the horizon increased,
346 eventually becoming the best performing forecast at four week horizons (rCRPS = 0.74) .

347 When forecasting infection incidence estimated from UK prevalence survey data[27], the no-
348 interaction model performed second only to the no-contact-data model (rCRPS = 0.89) for
349 horizons of one week, followed closely by the CoMix-data model (rCRPS=1.05). The POLYMOD-
350 data model performed worst when forecasting one week horizons with a rCRPS of 1.21.

351 However, at two week horizons the non-interaction model became the worst performing model
352 overall - which remained true for three and four week forecast horizons. In these cases the
353 CoMix-model performed best of all the models including the baseline models with rCRPS of 0.68,
354 0.64 and 0.57 (relative to the no-interaction model) for two, three and four week horizons
355 respectively. The second best performing NGM model at two and three week horizons was the
356 no-contact-data model, rCRPS of 0.82, 0.85 respectively. At four week horizons the POLYMOD-
357 data model was second best performing NGM model with a rCRPS of 0.76. The baselines both
358 did worse than all but the POLYMOD-data model when forecasting at a one week horizon,
359 however the performance of the fixed value baseline improved relative to all of the NGM models
360 at longer forecast horizons and produced the second best performing forecasts overall for
361 forecast horizons of three and four weeks (after the CoMix-data model) with rCRPS of 0.79 and
362 0.68 respectively.

363



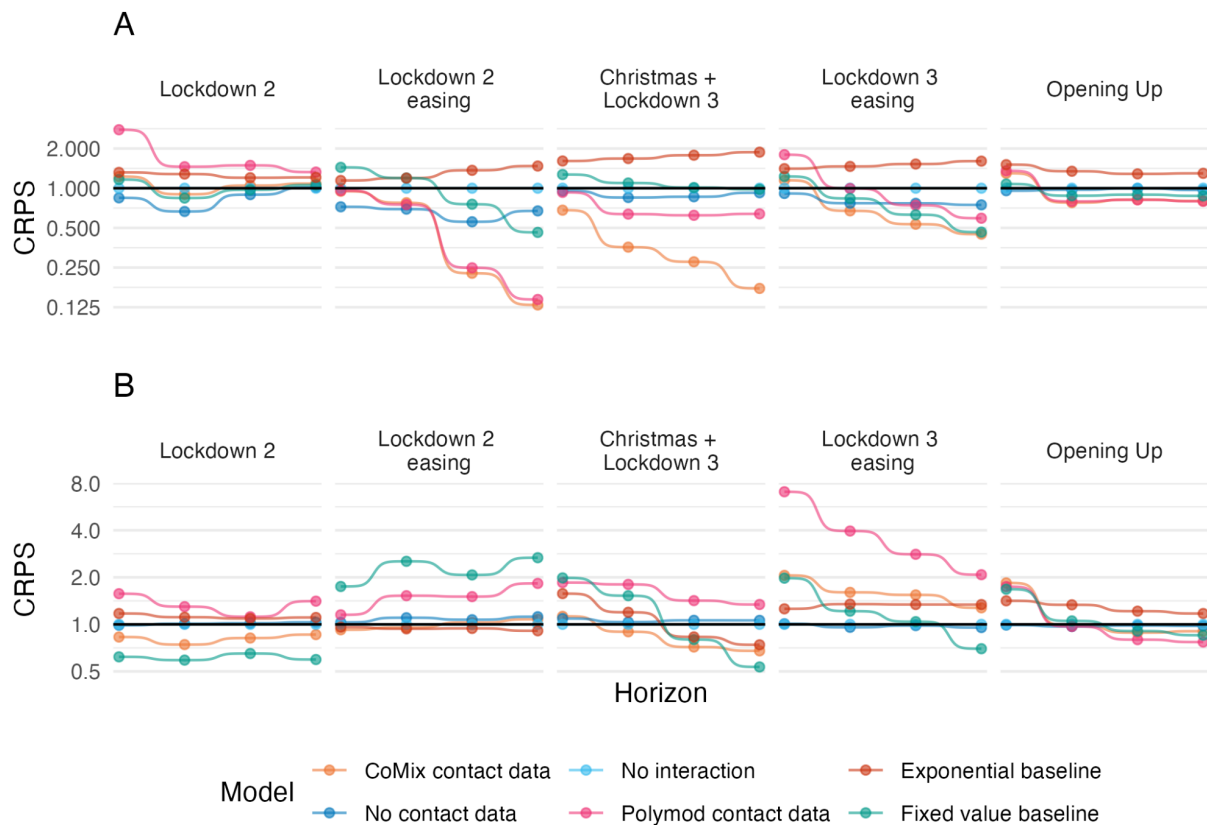
364

365 **Fig. 3** continuous ranked probability score relative to the score of the no-interaction model. A. shows
 366 overall performance of each model when applied to case (left) and infection (right) data with rCRPS
 367 on the y-axis and Bias on the x-axis. Forecast horizon is indicated by marker size. B. and C. show the
 368 CRPS relative to the no-interaction model against forecast horizon disaggregated by age for infection
 369 and case data respectively. The colour of the points shows the corresponding model.

370 We compared the relative forecast performance scoring predictions in each age-group
 371 separately (Figure 3 B and C, Table S2). When forecasting infection incidence, we found
 372 that the CoMix model and no-contact-data model forecast infections better than the no-
 373 interaction model in middle-aged adults and older adults (35+ years old) for all forecast
 374 horizons. The models also performed best at forecast horizons of two weeks or more in
 375 young children (2-10 years old) and older adults. In contrast, the non-interaction model
 376 performed much more similarly to the interaction models for forecasts within younger adults
 377 (16-34 years old). The same was also true for older age groups (60+) in the case forecasts
 378 but not for children, middle aged adults or children. For infections, the performance of all the
 379 models improved in all age categories relative to the exponential extrapolation baseline as
 380 forecast horizon increased, the fixed value baseline improved relative to the non-interaction

381 model in all age categories but provided poorer forecasts than the CoMix model in all age-
382 categories and time horizons. For case forecasts however the fixed value baseline improved
383 relative to all of the models as horizon increased, providing the best forecasts at four week
384 horizons in age groups between 0 and 59 years.

385 We divided the analysis into seven periods (Figure 4, Table S3), within each of which
386 national restrictions on social activity remained broadly consistent. For consistency we used
387 the same periods as those presented in Gimma et. al.[19], which presents the key findings of
388 the CoMix study. The relative improvement in performance for the CoMix-data model was
389 most consistent when forecasting infections during the *Christmas and Lockdown 3* period,
390 which was the only period where the CoMix-data model performed the best overall at all
391 forecast horizons. When forecasting cases, the CoMix-data model also performed relatively
392 well during this period at forecast horizons of two or more weeks, but performed comparably
393 to the no-interaction model at one week forecast horizon. The CoMix model's infection
394 forecasts outperformed all other NGM models in the two periods following this (*Lockdown 3*
395 *easing* and *Opening up*) for forecast horizons of two weeks or more, where only the fixed
396 value baseline model improved on its score. Similarly for the *Lockdown 2 easing* period, the
397 CoMix-data model performed better than the no-interaction model at all forecast horizons,
398 but the no-contact-data model performed better for one and two week forecast horizons. The
399 improved performance of the CoMix model was not wholly reflected in the case forecasts. In
400 particular, the CoMix model performed more poorly than the no-interaction model at all time
401 horizons during the *Lockdown 3 easing* period. However the CoMix model performed better
402 during the *Lockdown 2* period, than the other NGM models for case forecasts, whereas for
403 infection forecasts the CoMix model performed comparably to the no-interaction model
404 during this period.



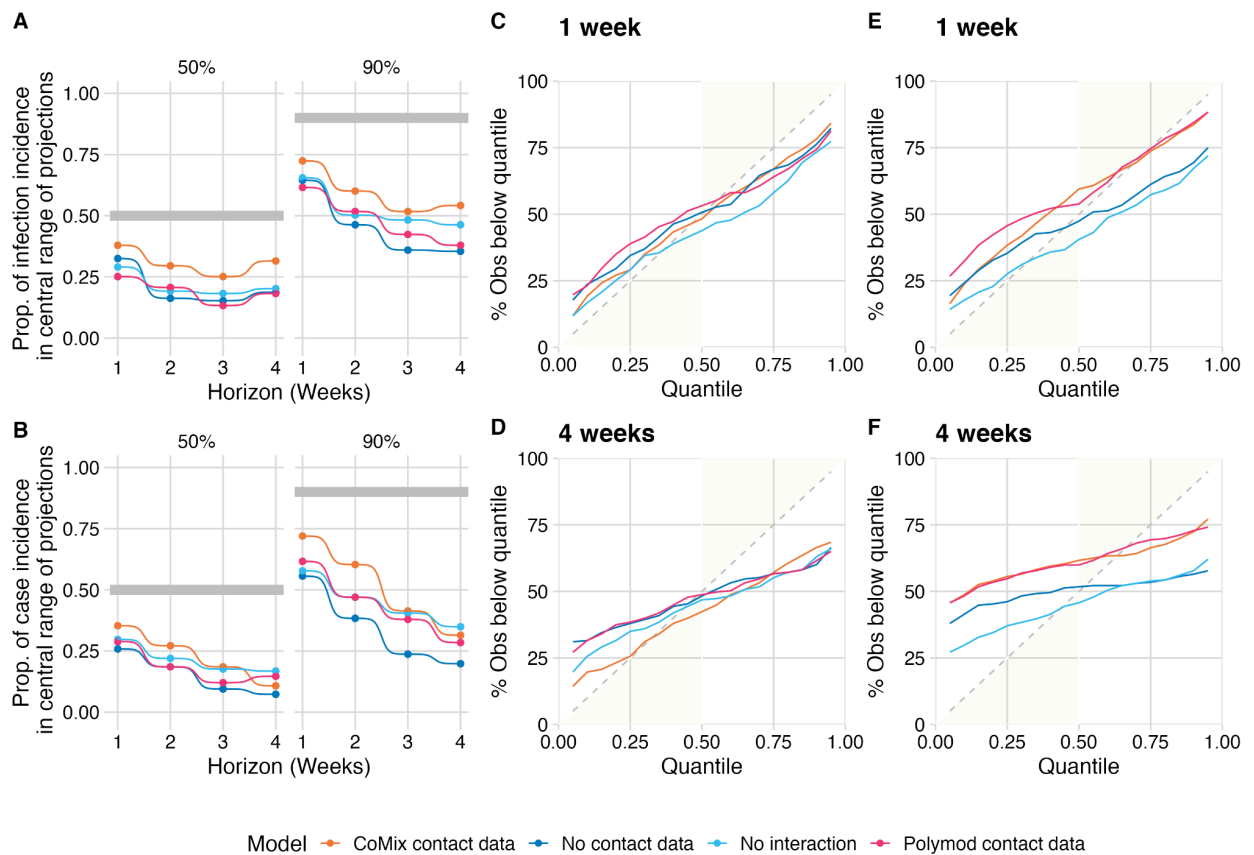
405

406 **Fig. 4** Continuous ranked probability score relative to the score of the no-interaction model against
 407 increasing forecast horizon (one to four weeks). Panels left to right show each pandemic period,
 408 Panels top to bottom show forecasts of cases and infections.

409 Forecast calibration

410 Of the four NGM models, the CoMix data-based model was best calibrated for case
 411 forecasts at one and two week horizons and at all horizons for infection forecasts. This is
 412 evidenced by closer agreement between the proportion of true values in each central range
 413 of the predictive distribution (50% and 90%) with the value of the range (Figure 5 A and B).
 414 Calibration typically became poorer at longer forecast horizons, with more true values falling
 415 outside the specified ranges than would be expected. We also note that none of the
 416 forecasts were particularly well calibrated when considered overall forecast dates. The best
 417 performing forecast, infection forecasts made by the CoMix model at a one week horizon,
 418 saw fewer than 75% of true values fall within the 90% confidence range of the associated

419 projections and fewer than 40% within the 75% confidence range. Separating the forecasts
420 by period of the pandemic (Supplementary Figures S7 and S8) revealed that the CoMix
421 model was best calibrated for ‘Christmas and Lockdown 3’ and ‘Lockdown 3’ periods, for
422 both the case and infection forecasts. In particular the CoMix model’s forecast of infections
423 was very well calibrated during the ‘Christmas and Lockdown 3’ period, with more than 80%
424 of true values falling within the 90% confidence range of the forecast at all horizons. The
425 other models were also relatively well calibrated during these periods. Overall the other
426 periods were much more poorly calibrated. In particular, the “Lockdown 2” and “Lockdown 2
427 easing” periods were very poorly calibrated across all models with no true values falling
428 within the 90% confidence range for the No-contact-data and No-interaction model forecasts
429 during the ‘Lockdown 2’ period. The baseline models are not presented as they present no
430 confidence ranges - hence the forecast coverage is zero for all forecasts by definition.



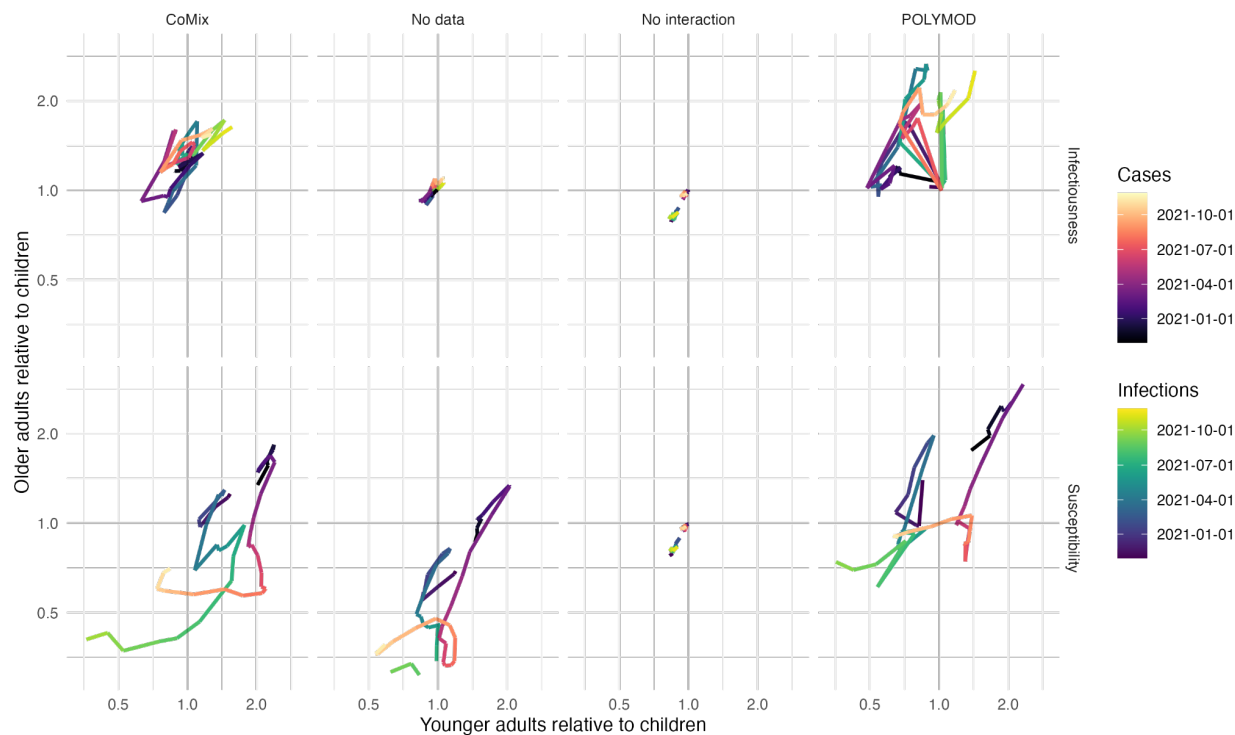
431

432 **Fig. 5** Calibration of the forecasts made by each of the next-generation-matrix-based models. A) The
 433 proportion of observed mean incidence of infection estimates (inc2prev) and B) the proportion of
 434 observed case numbers that fall within the 50% and 90% central interval of the relevant forecasts of
 435 the four models. C) and D) the percentage of observed mean incidence of infection estimates
 436 (inc2prev) falling below each quantile of the forecasts at one and four week horizons respectively. E)
 437 and F) the percentage of observed case counts falling below each quantile of the forecasts at one and
 438 four week horizons respectively.

439 Age-specific susceptibility and infectiousness

440 We extracted estimates of age specific infectiousness and susceptibility from the model fits
 441 to assess the biological plausibility of these parameters (Figure 6). For the CoMix model the
 442 susceptibility of younger adults (16-49 years old for infections and 20-49 for cases) and older
 443 adults (50+ years old) was higher relative to children (under 16 for infections and under 20
 444 for cases). In the early part of 2021, this began to shift such that first susceptibility reduced

445 relative to children in the older adults and then in younger adults. Ultimately by the end of the
446 period evaluated, children had higher susceptibility relative to all adults. A similar pattern
447 was present in all models that allowed interaction between age-groups. Infectiousness
448 broadly remained equal between age-groups, with the exception of a small number of
449 outlying values within which there is no clear trend.



450

451 **Fig. 6** Infectiousness (top panels) and susceptibility (bottom panels) of younger adults (16-50 for
452 infections and 20-50 for cases) against that of older adults (50+) each relative to children (under 16
453 for infections and under 20 for cases) by forecast date (hue). Each model is shown in panels left to
454 right.

455

456

457 Discussion

458 Evaluating the forecast performance of four next generation matrix models, we found that
459 allowing interaction between age-groups and using regularly collected contact data did not
460 consistently improve performance when forecasting cases as reported through the UK
461 COVID-19 dashboard. However, when forecasting infection incidence estimated from
462 prevalence survey data, forecast performance was improved overall by allowing interaction
463 between age-groups at all time horizons. We found that informing interaction by regularly
464 collected contact data improved forecasts further at time horizons of two weeks and greater.
465 Although we found that the improvement was not consistent across all periods or when
466 considering the resulting forecasts for each age-group separately.

467 The NGM models with interaction showed the most benefit over the no-interaction model
468 when forecasting infections during the *Christmas and Lockdown 3* period. Here the CoMix-
469 data model outperformed all other models. During this period the CoMix-data model also
470 proved to be the best calibrated of any model during any period. It's notable that the
471 forecasts being made at a time where the most intense restrictions were imposed for an
472 extended period of time following a very sharp rise in cases. The sharp rise in cases
473 combined with growing hospitalisations and deaths may have resulted in a period of
474 consistent behaviour amongst the population, since although restrictions changed on
475 January 5th[42] the contact behaviour recorded by CoMix remained similar between
476 *Christmas and Lockdown 3*[19]. This consistency of behaviour, well described by CoMix
477 data, over an extended period of time may ultimately support the performance of this model
478 over the others.

479 Generally, the NGM models that allowed interaction between age-groups performed better
480 than the model with no interaction for infection forecasts. This was particularly true when
481 considering performance in older and younger age-groups. This effect may relate to the age-
482 specific incidence and transmission rates. Whereas infections in the younger adults groups

483 were largely driven by transmission within the age-group, for long periods of the pandemic
484 infections in elderly and children are likely to have been driven primarily by transmission
485 from the younger adults age-groups, particularly when schools were closed, which was the
486 case for a large proportion of the studied period[7,43]—making incidence projections in
487 these groups more reliant on between-age-group interaction.

488 Overall, all forecasts performed better than the exponential extrapolation baseline when
489 forecasting infections. The relative performance of this baseline compared to all other
490 models generally worsened over time suggesting that the simplistic exponential growth
491 model tends to overestimate any change in infections over time - which is compounded at
492 longer horizons. Although the relative performance of the simple exponential extrapolation
493 was better for case forecasts than infection forecasts at short time horizons, similarly to the
494 infection forecasts, all models improved relative to this baseline as the forecast horizon
495 increased, mostly surpassing it to provide better predictions at a longer horizon, showing that
496 this simple assumption of transmission dynamics breaks down rapidly.

497 In contrast, in both case and infection forecasts, the relative performance of the fixed value
498 baseline improved with increased forecast horizon as all of the modelled values deviated
499 from the true values over time in the case of all models. This may represent the rapidly
500 changing behaviour of the public under constantly changing interventions. This however,
501 also compounds existing evidence that effective forecasts of infectious disease incidence
502 can rarely be made at horizons of greater than a few weeks[6].

503 The distributions of relative infectiousness and susceptibility inferred by the models are
504 consistent with others findings, beginning with adults exhibiting higher susceptibility than
505 children in general[7,8,16]. This changes throughout the pandemic, following a sequence
506 consistent with what may be expected as a result of acquisition of antibodies through
507 vaccination and natural infection. The largest changes occur after vaccination is introduced,
508 where the susceptibility of the older adults reduces relative to other ages first, followed by

509 susceptibility of younger adults. This is consistent with the vaccine roll out schedule in
510 England during the early part of 2021[44]. The general trajectory of age-specific
511 susceptibility also agrees well with findings of Franco et. al.[44,45] which used the Belgian
512 arm of the CoMix study to estimate age-specific infectiousness and susceptibility
513 independently to this study.

514 Our estimates of age-specific infectiousness and susceptibility need to be interpreted with
515 caution for three main reasons. Firstly, the framework is optimised for prediction as opposed
516 to inference and therefore is not set up to best reflect the biological processes at play but
517 rather to make good predictions. Secondly, there is likely to be some bias in the way contact
518 data is collected by age which may impact these estimates[19]. Importantly, contacts of
519 children are reported by their parents or guardians. In addition, children's contacts are
520 disproportionately reported as groups—markedly different from adult contacts, which were
521 reported by the participant themselves and were mostly reported as individual contacts.
522 Finally, we also make no differentiation between contacts by location, duration or intimacy. In
523 reality contacts made in different contexts (e.g. home and school) are likely to carry different
524 potential of transmission, which may also affect the way our susceptibility and infectiousness
525 estimates can be interpreted. One potential extension would be to include contacts by
526 setting (Work, School, Home and Other), which would allow contacts in different contexts to
527 be weighted differently.

528 There may also be other factors associated with inferred changes in susceptibility and
529 infectiousness which do not correspond to inherent transmissibility. For example, the degree
530 of mitigating behaviours unrelated to contact rate (e.g., masks, preferring outdoor meetings,
531 physical distancing) may have changed differently over the epidemic for each age group. A
532 reduction in relative susceptibility in older adults may indicate that these age groups were
533 able to reduce the risk of infection even when making contact with others further into the
534 pandemic than younger age groups. Also, we assumed immunity is determined by
535 seropositivity as reported in the publicly available CIS data[27], from which we only used a

536 single antibody level threshold for positivity. It may be the case that there is substantial
537 variation in the antibody level distribution in sero-positive individuals of different age groups
538 based on the distribution of vaccine history and infection acquired antibodies, which may
539 affect age-stratified susceptibility to infection. Finally, there may be variation in the age-
540 profile of susceptibility by variant, however due to the limitations discussed, we are unable to
541 quantify this here.

542 Whereas the relative performance of the models was fairly consistent for infections, the
543 performance when applied to case data was generally more erratic with the ranking for
544 models and baselines changing between horizons within the same aggregation of forecast
545 dates and age groups. This may reflect the more variable nature of case reporting, which is
546 affected by multiple external factors affecting the week-by-week variation in cases beyond
547 transmission dynamics. Notably, case reports are subject to variation in reporting rate, which
548 may also differ between age-groups. This is exacerbated by changes in the UK
549 Government's testing policy over the course of the pandemic. This was not the case with the
550 infection time-series, which was estimated from weekly prevalence estimates. Moreover, the
551 infection forecasts incorporated estimates of antibody prevalence modelled from weekly
552 serosurveys[27] and vaccination data[29], whereas the case forecasts did not.

553 Our work provides an indication of the potential benefits of including contact data in
554 epidemiological forecasts, but for transparency in our analysis we have chosen not to use
555 state of the art methods of surveillance, instead there are a number of simplifications we
556 made when selecting and processing the epidemiological data to provide clearer analysis of
557 this effect. These simplifications would be expected to affect the performance of forecasts
558 when implemented in real time. Firstly, in our analysis we forecast infections using a
559 modelled time-series fit to weekly prevalence estimates[26]. In truth, under the current data
560 sharing protocol of the ONS Covid-19 infection survey, this data would not be publicly
561 available on the forecast date and hence is not, in this form, applicable as a real-time
562 application without fully integrating into the ONS infection survey workflow. We chose to do

563 this to provide the most idealised scenario to test the application of contact data to short
564 term forecasts, without the complexities associated with case data. Furthermore, although
565 these estimates agree well with other estimates and case time-series, the methodology
566 promotes a smooth infection history leading to autocorrelation in the time-series. This may
567 unduly benefit models with high autocorrelation properties, e.g., the fixed value baseline.
568 However, the similar relative performance of this model when evaluating case data supports
569 our observations that this model performs best at longer time horizons. Secondly, an
570 important feature of real-time epidemiological data is that there are several complex delay
571 distributions which may affect the recent time series of cases[46]. This is especially true
572 when using data by date of specimen as we do here, where full information of cases at
573 specimen date are not available until all tests from that date have been processed. For this
574 reason, case counts are increasingly truncated in the days leading up to the forecast date.
575 Approaches to account for this exist[28], but here we have used the case data as known
576 now as opposed to as known on each forecast date, as such we did not need to make this
577 adjustment—as we would if we were making the forecasts in real-time. Extending existing
578 approaches for real-time modelling that can deal with truncated data to include interactions
579 between multiple time series will be an important area of future research[47,48].

580 The models we present used a normal likelihood, unconventional for epidemiological
581 forecasts which tend to operate on count data. In our case, we use estimates of infection
582 incidence, our input data is therefore not an integer time series, but a distribution at each
583 time-point. To keep the estimates consistent between the case and infection time series' we
584 maintained this approach for forecasting cases as well. Lastly, the absolute measure
585 provided by CRPS means that the overall score is weighted towards age-groups and time
586 periods where the absolute incidence was high, this may negatively impact the overall score
587 of models which did poorly in the “young adults” age range (16-35) where incidence was
588 highest for much of the study period.

589 Overall, allowing interaction between age-groups and integrating regularly collected contact
590 data improved forecasts when forecasting infections based on estimates from national
591 prevalence surveys. This benefit was, however, not clear when applied to regularly collected
592 case data, which is generally much more readily available for real-time applications. The
593 picture this offers of the usefulness of contacts in forecasts is nuanced. Even for the
594 idealised example of incident infections estimated retrospectively from repeated cross-
595 sectional prevalence surveys, there are periods of improved performance, and times where
596 the contact-based models failed to capture the dynamics of the epidemic sufficiently to
597 improve on the other models' predictions. The period where the contact data performed the
598 best was during a period where contacts remained relatively consistent. This raises the
599 question as to whether real-time contact data is capable of capturing relevant changes in
600 transmission related behaviour when implementation of non-pharmaceutical interventions
601 are regularly changed. As applications using contact data in real-time develop, it is important
602 to evaluate whether the periods where contact data are informative are aligned with periods
603 when they are also useful for infection control, and consider how future studies might be
604 optimised to ensure this target can be achieved.

605 **Acknowledgements**

606 The authors would like to acknowledge the contributions of colleagues from the COVID-19
607 infection Survey Analysis team at the Office for National Statistics (ONS) for their project
608 support and thoughtful discussion during the planning and analysis phase of this research.
609 Also, the members of the CoMix consortium for their support with the contact data,
610 especially Chris Jarvis, Pietro Colletti, Niel Hens and John Edmunds for their feedback on
611 the manuscript. Thirdly, Lloyd Chapman for insightful discussion during the analysis phase of
612 the work. Finally, members of the Epiforecasts group at LSHTM for helpful comments and
613 feedback on our modelling framework, especially Nikos Bosse for support with the
614 *scoringutils* package.

615 References

- 616 1. Thacker SB, Choi K, Brachman PS. The surveillance of infectious diseases. *JAMA*.
617 1983;249: 1181–1185.
- 618 2. Fraser C. Estimating individual and household reproduction numbers in an emerging
619 epidemic. *PLoS One*. 2007;2: e758.
- 620 3. Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to
621 estimate time-varying reproduction numbers during epidemics. *Am J Epidemiol*.
622 2013;178: 1505–1512.
- 623 4. Funk S, Camacho A, Kucharski AJ, Eggo RM, Edmunds WJ. Real-time forecasting of
624 infectious disease dynamics with a stochastic semi-mechanistic model. *Epidemics*.
625 2018;22: 56–61.
- 626 5. Sherratt K, Gruson H, Grah R, Johnson H, Niehus R, Prasse B, et al. Predictive
627 performance of multi-model ensemble forecasts of COVID-19 across European nations.
628 medRxiv. 2022. doi:10.1101/2022.06.16.22276024
- 629 6. Cramer EY, Ray EL, Lopez VK, Bracher J, Brennen A, Castro Rivadeneira AJ, et al.
630 Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in
631 the United States. *Proc Natl Acad Sci U S A*. 2022;119: e2113561119.
- 632 7. Davies NG, Klepac P, Liu Y, Prem K, Jit M, CMMID COVID-19 working group, et al.
633 Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat*
634 *Med*. 2020;26: 1205–1211.
- 635 8. Viner RM, Mytton OT, Bonell C, Melendez-Torres GJ, Ward J, Hudson L, et al.
636 Susceptibility to SARS-CoV-2 Infection Among Children and Adolescents Compared
637 With Adults: A Systematic Review and Meta-analysis. *JAMA Pediatr*. 2021;175: 143–
638 156.
- 639 9. COVID-19 Forecasting Team. Variation in the COVID-19 infection-fatality ratio by age,
640 time, and geography during the pre-vaccine era: a systematic analysis. *Lancet*.
641 2022;399: 1469–1488.
- 642 10. Levin AT, Hanage WP, Owusu-Boaitey N, Cochran KB, Walsh SP, Meyerowitz-Katz G.
643 Assessing the age specificity of infection fatality rates for COVID-19: systematic review,
644 meta-analysis, and public policy implications. *Eur J Epidemiol*. 2020;35: 1123–1138.
- 645 11. Pijls BG, Jolani S, Atherley A, Derckx RT, Dijkstra JIR, Franssen GH, et al.
646 Demographic risk factors for COVID-19 infection, severity, ICU admission and death: a
647 meta-analysis of 59 studies. *BMJ Open*. 2021;11: e044640.
- 648 12. Anderson RM, May RM. Age-related changes in the rate of disease transmission:
649 implications for the design of vaccination programmes. *J Hyg* . 1985;94: 365–436.
- 650 13. Wallinga J, Teunis P, Kretzschmar M. Using data on social contacts to estimate age-
651 specific transmission parameters for respiratory-spread infectious agents. *Am J*
652 *Epidemiol*. 2006;164: 936–944.
- 653 14. Worby CJ, Chaves SS, Wallinga J, Lipsitch M, Finelli L, Goldstein E. On the relative role
654 of different age groups in influenza epidemics. *Epidemics*. 2015;13: 10–16.

- 655 15. Hoang T, Coletti P, Melegaro A, Wallinga J, Grijalva CG, Edmunds JW, et al. A
656 Systematic Review of Social Contact Surveys to Inform Transmission Models of Close-
657 contact Infections. *Epidemiology*. 2019;30: 723–736.
- 658 16. House T, Riley H, Pellis L, Pouwels KB, Bacon S, Eidukas A, et al. Inferring Risks of
659 Coronavirus Transmission from Community Household Data. *arXiv e-prints*. 2021;
660 arXiv:2104.04605.
- 661 17. Edmunds WJ, O’Callaghan CJ, Nokes DJ. Who mixes with whom? A method to
662 determine the contact patterns of adults that may lead to the spread of airborne
663 infections. *Proc Biol Sci*. 1997;264: 949–957.
- 664 18. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social contacts
665 and mixing patterns relevant to the spread of infectious diseases. *PLoS Med*. 2008;5:
666 e74.
- 667 19. Gimma A, Munday JD, Wong KLM, Coletti P, van Zandvoort K, Prem K, et al. CoMix:
668 Changes in social contacts as measured by the contact survey during the COVID-19
669 pandemic in England between March 2020 and March 2021. *bioRxiv*. medRxiv; 2021.
670 doi:10.1101/2021.05.28.21257973
- 671 20. Munday JD, Jarvis CI, Gimma A, Wong KLM, van Zandvoort K, CMMID COVID-19
672 Working Group, et al. Estimating the impact of reopening schools on the reproduction
673 number of SARS-CoV-2 in England, using weekly contact survey data. *BMC Med*.
674 2021;19: 233.
- 675 21. Jarvis CI, Gimma A, van Zandvoort K, Wong KLM, CMMID COVID-19 working group,
676 Edmunds WJ. The impact of local and national restrictions in response to COVID-19 on
677 social contacts in England: a longitudinal natural experiment. *BMC Med*. 2021;19: 52.
- 678 22. Verelst F, Hermans L, Vercruyse S, Gimma A, Coletti P, Backer JA, et al. SOCRATES-
679 CoMix: a platform for timely and open-source contact mixing data during and in between
680 COVID-19 surges and interventions in over 20 European countries. *BMC Med*. 2021;19.
681 doi:10.1186/s12916-021-02133-y
- 682 23. Funk S, Camacho A, Kucharski AJ, Lowe R, Eggo RM, Edmunds WJ. Assessing the
683 performance of real-time epidemic forecasts: A case study of Ebola in the Western Area
684 region of Sierra Leone, 2014–15. *PLoS Comput Biol*. 2019;15: e1006785.
- 685 24. Gneiting T, Raftery AE. Strictly Proper Scoring Rules, Prediction, and Estimation. *J Am*
686 *Stat Assoc*. 2007;102: 359–378.
- 687 25. Held L, Meyer S, Bracher J. Probabilistic forecasting in infectious disease epidemiology:
688 the 13th Armitage lecture. *Stat Med*. 2017;36: 3443–3460.
- 689 26. Abbott S, Funk S. Estimating epidemiological quantities from repeated cross-sectional
690 prevalence measurements. *bioRxiv*. 2022. doi:10.1101/2022.03.29.22273101
- 691 27. Coronavirus (COVID-19) Infection Survey, UK Statistical bulletins. [cited 29 Mar 2022].
692 Available:
693 [https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditions](https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/previousReleases)
694 [anddiseases/bulletins/coronaviruscovid19infectionsurveypilot/previousReleases](https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/previousReleases)
- 695 28. Bastos LS, Economou T, Gomes MFC, Villela DAM, Coelho FC, Cruz OG, et al. A
696 modelling approach for correcting reporting delays in disease surveillance data. *Stat*
697 *Med*. 2019;38: 4363–4377.

- 698 29. HM Government UK. Official Coronavirus (COVID-19) disease situation dashboard.
699 [cited 25 May 2022]. Available: <https://coronavirus.data.gov.uk/>
- 700 30. Uk G. Number of coronavirus (COVID-19) cases and risk in the UK. online: <https://www.gov.uk/guidance/coronavirus-covid-19-information-for-the-public>. 2020. Available:
701 [https://www.bleadon.org.uk/media/other/24400/NumberofcoronavirusCOVID-](https://www.bleadon.org.uk/media/other/24400/NumberofcoronavirusCOVID-19casesandriskintheUK-GOV.UK.pdf)
702 [19casesandriskintheUK-GOV.UK.pdf](https://www.bleadon.org.uk/media/other/24400/NumberofcoronavirusCOVID-19casesandriskintheUK-GOV.UK.pdf)
703
- 704 31. inc2prev: Estimate incidence from ONS prevalence estimates. Github; Available:
705 <https://github.com/epiforecasts/inc2prev>
- 706 32. The CoMix study. In: uHasselt [Internet]. [cited 29 Mar 2022]. Available:
707 <https://www.uhasselt.be/en/aparte-sites-partner-en/epipose/the-comix-study>
- 708 33. Stan Team. Stan Modeling Language User's Guide and Reference Manual. 2012.
- 709 34. Madewell ZJ, Yang Y, Longini IM Jr, Halloran ME, Dean NE. Household Secondary
710 Attack Rates of SARS-CoV-2 by Variant and Vaccination Status: An Updated
711 Systematic Review and Meta-analysis. *JAMA Netw Open*. 2022;5: e229317.
- 712 35. Ganyani T, Kremer C, Chen D, Torneri A, Faes C, Wallinga J, et al. Estimating the
713 generation interval for coronavirus disease (COVID-19) based on symptom onset data,
714 March 2020. *Euro Surveill*. 2020;25. doi:10.2807/1560-7917.ES.2020.25.17.2000257
- 715 36. Hart WS, Miller E, Andrews NJ, Waight P, Maini PK, Funk S, et al. Generation time of
716 the alpha and delta SARS-CoV-2 variants: an epidemiological analysis. *Lancet Infect
717 Dis*. 2022;22: 603–610.
- 718 37. Alene M, Yismaw L, Assemie MA, Ketema DB, Gietaneh W, Birhan TY. Serial interval
719 and incubation period of COVID-19: a systematic review and meta-analysis. *BMC Infect
720 Dis*. 2021;21: 257.
- 721 38. Abbott S, Sherratt K, Gerstung M, Funk S. Estimation of the test to test distribution as a
722 proxy for generation interval distribution for the Omicron variant in England. *bioRxiv*.
723 2022. doi:10.1101/2022.01.08.22268920
- 724 39. McAloon C, Collins Á, Hunt K, Barber A, Byrne AW, Butler F, et al. Incubation period of
725 COVID-19: a rapid systematic review and meta-analysis of observational research. *BMJ
726 Open*. 2020;10: e039652.
- 727 40. Report 49 - Growth, population distribution and immune escape of Omicron in England.
728 In: Imperial College London [Internet]. [cited 29 Mar 2022]. Available:
729 [https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-49-](https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-49-Omicron/)
730 [Omicron/](https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-49-Omicron/)
- 731 41. Bosse NI, Abbott S, EpiForecasts FS. scoringutils: Utilities for Scoring and Assessing
732 Predictions. 2020.
- 733 42. Institute for Government. Timeline of UK government coronavirus lockdowns and
734 restrictions. [cited 28 Jul 2022]. Available:
735 <https://www.instituteforgovernment.org.uk/charts/uk-government-coronavirus-lockdowns>
- 736 43. Monod M, Blenkinsop A, Xi X, Hebert D, Bershan S, Tietze S, et al. Age groups that
737 sustain resurging COVID-19 epidemics in the United States. *Science*. 2021;371.
738 doi:10.1126/science.abe8372
- 739 44. COVID-19 vaccination programme. In: GOV.UK [Internet]. 27 Nov 2020 [cited 28 Jul

- 740 2022]. Available: [https://www.gov.uk/government/collections/covid-19-vaccination-](https://www.gov.uk/government/collections/covid-19-vaccination-programme)
741 programme
- 742 45. Franco N, Coletti P, Willem L, Angeli L, Lajot A, Abrams S, et al. Inferring age-specific
743 differences in susceptibility to and infectiousness upon SARS-CoV-2 infection based on
744 Belgian social contact data. PLoS Comput Biol. 2022;18: e1009965.
- 745 46. Gostic KM, McGough L, Baskerville EB, Abbott S, Joshi K, Tedijanto C, et al. Practical
746 considerations for measuring the effective reproductive number, Rt. PLoS Comput Biol.
747 2020;16: e1008409.
- 748 47. Abbott S, Lison A, Funk S. epinowcast: Flexible hierarchical nowcasting. Zenodo; 2022.
749 doi:10.5281/ZENODO.5637165
- 750 48. Abbott S, Hellewell J, Sherratt K, Gostic K, Hickson J, Badr HS, et al. EpiNow2:
751 Estimate Real-Time Case Counts and Time-Varying Epidemiological Parameters. 2021.
752 doi:10.5281/ZENODO.3957489
- 753
- 754
- 755

756 **Author Contributions**

757 JDM and SF Were responsible for funding acquisition, conceptualization and methodology,
758 data curation and formal analysis of the infection estimates. JDM was responsible for project
759 administration, carried out the Formal analysis and investigation of forecasts and prepared
760 the original draft of the manuscript and visualisations with supervision from SF. JDM, SF, SA
761 and SM reviewed and edited the manuscript.

762 **Data availability statement**

763 Case data is available on the UK Covid-19 Dashboard <https://coronavirus.data.gov.uk>
764 Infection and antibody data are available on the Covid-19 infection survey website
765 <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveyspilot/25november2022> and contact
766 matrices are available from the CoMix online repository
767 <https://doi.org/10.5281/zenodo.7351951>.

769 **Code availability statement**

770 All code used for this work is available at
771 <https://github.com/epiforecasts/CovidAgeGroupForecast>. The weekly contact matrices are
772 published on Zenodo <https://doi.org/10.5281/zenodo.7351951>

773 **Financial Disclosure Statement**

774 This work was partly funded by an Office for National Statistics COVID-19 Infection Survey
775 Analysis grant PU-20-0205(c): JDM. This work was partly funded by the Wellcome Trust
776 210758/Z/18/Z: JDM and SF

777 **Competing interests**

778 The authors declare that they have no competing interests