

1 **Nowcasting and Forecasting the 2022 U.S. Mpox Outbreak: Support for Public Health**

2 **Decision Making and Lessons Learned**

3 Kelly Charniga, PhD^{1,2*}, Zachary J. Madewell, PhD^{1,3}, Nina B. Masters, PhD⁴, Jason Asher, PhD⁵,

4 Yoshinori Nakazawa, PhD², Ian H. Spicknall, PhD⁶

5

6 ¹ These first authors contributed equally to this article.

7 ² Division of High Consequence Pathogens and Pathology, National Center for Emerging and Zoonotic
8 Infectious Diseases, CDC

9 ³ Center for Global Health, CDC

10 ⁴ Epidemic Intelligence Service, CDC

11 ⁵ Center for Forecasting and Outbreak Analytics, CDC

12 ⁶ Division of Sexually Transmitted Disease Prevention, National Center for HIV, Viral Hepatitis, STD, &
13 TB Prevention, CDC

14

15 * Corresponding author, email: ruq7@cdc.gov

16

17

18

19

20

21

22

23

24

25

26 Keywords: monkeypox virus, modeling, epidemic, real-time analysis, situational awareness

27

28 **Abstract**

29 In June of 2022, the U.S. Centers for Disease Control and Prevention (CDC) Mpox Response wanted
30 timely answers to important epidemiological questions which can now be answered more effectively
31 through infectious disease modeling. Infectious disease models have shown to be valuable tool for
32 decision making during outbreaks; however, model complexity often makes communicating the results
33 and limitations of models to decision makers difficult. We performed nowcasting and forecasting for the
34 2022 mpox outbreak in the United States using the R package EpiNow2. We generated
35 nowcasts/forecasts at the national level, by Census region, and for jurisdictions reporting the greatest
36 number of mpox cases. Modeling results were shared for situational awareness within the CDC Mpox
37 Response and publicly on the CDC website. We retrospectively evaluated forecast predictions at four key
38 phases during the outbreak using three metrics, the weighted interval score, mean absolute error, and
39 prediction interval coverage. We compared the performance of EpiNow2 with a naïve Bayesian
40 generalized linear model (GLM). The EpiNow2 model had less probabilistic error than the GLM during
41 every outbreak phase except for the early phase. We share our experiences with an existing tool for
42 nowcasting/forecasting and highlight areas of improvement for the development of future tools. We also
43 reflect on lessons learned regarding data quality issues and adapting modeling results for different
44 audiences.

45 **Background**

46 The 2022 mpox (formerly known as monkeypox) outbreak is the first major infectious disease
47 outbreak since the COVID-19 pandemic and was declared a Public Health Emergency of International
48 Concern by the World Health Organization on July 23, 2022 (1). As of April 13, 2023, a total of 86,956
49 confirmed cases have been reported in 110 countries and territories (2). Unlike COVID-19, mpox is a
50 disease known to be endemic in West and Central Africa for decades; it is caused by monkeypox virus
51 (MPXV), a zoonotic orthopoxvirus (3). Historically, classical symptoms involved fever, headache, muscle
52 aches, fatigue, lymphadenopathy, and rash (4). Human-to-human MPXV transmission occurs through
53 close contact with infectious material from skin lesions, respiratory secretions during prolonged face-to-
54 face contact, and fomites, such as linens and bedding (5). The 2022 mpox outbreak began in May and
55 spread rapidly in non-endemic countries. This outbreak was characterized by human-to-human
56 transmission of MPXV through close physical contact (often associated with sexual activities) and has
57 disproportionately affected gay, bisexual, and other men who have sex with men (6).

58 During a public health crisis such as the mpox outbreak, difficult and rapid decisions with limited
59 available data are often required (7). Infectious disease models may assist with informing policy and
60 practice by predicting the magnitude and duration of an outbreak or epidemic, evaluating characteristics
61 of pathogen transmission such as transmissibility, and designing vaccination strategies, among others (8,
62 9). However, infectious disease models are often complex, integrating data from heterogenous sources
63 with many parameter assumptions that are subject to uncertainty. These aspects make it challenging to
64 effectively implement such models and communicate the results and potential limitations to decision
65 makers, other public health partners, and the general public (10, 11).

66 During the COVID-19 pandemic, the state of the art of outbreak analysis advanced considerably
67 (12). Methods and tools for estimating key epidemiological parameters, such as the effective reproduction
68 number, R_t , were developed and shared in real-time (13). Monitoring R_t , the average number of secondary
69 cases caused by a single infected individual in a large population, during an outbreak is useful for
70 assessing transmission dynamics and evaluating the effectiveness of public health measures (e.g.,

71 vaccination, contact tracing, isolation, and quarantine) (14). Nowcasts and forecasts have been produced
72 by numerous research groups around the globe (15-18), the results of which were instrumental for
73 decision makers weighing possible control measures (19) such as social distancing measures. Outbreak
74 forecasting predicts specific outcomes (e.g., number of cases, deaths, or hospitalizations) at some specific
75 future times (e.g., weeks, months, etc.), whereas nowcasting estimates those outcomes for the current
76 time, accounting for delays in reporting.

77 In this manuscript, we share our experience nowcasting and forecasting the mpox outbreak,
78 including adapting the modeling output to different audiences. We also describe challenges faced vis-a-
79 vis data quality, parameter estimation, and model application and propose ways to improve nowcasting
80 and short-term forecasting efforts for future outbreaks.

81 **Methods**

82 **Nowcasting/forecasting the mpox outbreak**

83 We used data on probable and confirmed mpox cases in the United States (see “Case definition”
84 in Supplementary methods) reported to CDC by state and local public health jurisdictions from May 17,
85 2022, through March 16, 2023. Data were submitted in several different formats throughout the outbreak
86 period. These formats included: a CDC-operated call center through the Emergency Operations Center
87 (EOC), a long and a short case report form (CRF), and the National Notifiable Diseases Surveillance
88 System (NNDSS). Cases could have data submitted via more than one format and jurisdictions could
89 update data on cases after initial submission (Supplementary methods). All reported data were processed
90 in CDC’s Data Collation and Integration for Public Health Event Response (DCIPHER) platform, an
91 instance of Palantir Foundry (Palantir Technologies Inc, Denver, CO). DCIPHER is a secure, cloud-based
92 data integration, analytics, and situational awareness platform used by the Centers for Disease Control
93 and Prevention (CDC), federal partners, and state, tribal, local, and territorial public health jurisdictions to
94 collect, collaborate on, and share public health data (20). DCIPHER collates data of differing origin,
95 structure, and purpose to provide near real-time insights into public health problems, with the goal of
96 providing a complete picture of situational awareness.

97 We considered three approaches for estimating R_t which are implemented in the R packages
98 EpiEstim (version 2.2-4) (21), earlyR (version 0.0.5) (22), and EpiNow2 (version 1.3.2) (23)
99 (Supplementary methods). Initially, we used all three methods to estimate R_t at the national level as well
100 as for jurisdictions reporting the highest incidence of mpox. Although estimates of the historical range of
101 the serial interval of mpox were available at the start of the outbreak, they were based on data from the
102 Democratic Republic of Congo, which reflected largely non-sexual household spread (24). We considered
103 these historical parameter estimates as a starting point for early outbreak analysis, using them (along with
104 sensitivity analyses) until new estimates were generated. Updated estimates characterized by the mean
105 and standard deviation were needed for the global outbreak given the novel mode of transmission. In June
106 2022, we were able to use an estimated mean serial interval (i.e., the period of time between symptom
107 onset in the primary case and symptom onset in the secondary case) of 9.8 days (95% credible interval
108 [CrI]: 5.9 – 21.4) from 17 case pairs reported by the United Kingdom (6). At that time, symptom onset
109 date was available for most reported cases, and imported cases were still contributing to a high proportion
110 of MPXV transmission. EpiEstim results were considered the most appropriate at this stage of the
111 outbreak because this method accounts for imported vs. locally acquired cases, has a stable codebase, is
112 widely used, and is computationally efficient (25).

113 In July 2022, we started exclusively using EpiNow2, which uses a similar approach as EpiEstim
114 (a branching process model, Supplementary methods), but it better accounts for reporting delays and
115 incorporates multiple sources of uncertainty (13); for example, it removes noise associated with weekend
116 effects and uses random walks for temporal smoothing. Forecasting is supported internally for R_t , number
117 of infections, cases by date of report, and growth rate. Unlike EpiEstim, EpiNow2 does not distinguish
118 between imported vs. locally acquired infections. EpiNow2 is the most computationally expensive
119 approach, requiring longer model run times (Supplementary methods). The model assumes that testing
120 procedures, surveillance effort, and reporting delays remain constant over the estimation period. To use
121 EpiNow2, cases by date of report must be provided as well as the generation time distribution (the time
122 between infection of a primary and secondary case), incubation period distribution (the time between

123 infection and symptom onset in a case), and any other delay distributions (e.g., the delay between
124 symptom onset and report date). The model estimates the number of new cases by date of report, number
125 of cases by their date of infection, R_t , and time-varying growth rate. Estimates over the last 16 days of the
126 time-series are based on partial data due to the presumption of reporting delays. Input parameters and
127 methods for adjusting for right-truncation evolved as we learned more about the outbreak.

128 **Communication methods**

129 R_t estimates were shared internally through Situational Reports and leadership meetings and
130 publicly through CDC's Technical Reports (26) and CDC's public-facing mpox website (27). The
131 Technical Reports were co-led by the Center for Forecasting and Outbreak Analytics (CFA) and the 2022
132 Multi-National Mpox Outbreak Response. Estimates were generated for distribution at least once per
133 week. The technical reports were intended for scientific audiences. The purpose of sharing these results
134 was to improve understanding of the outbreak and inform further scientific inquiry.

135 **Performance assessment methods**

136 We chose eight dates during four key outbreak phases to retrospectively evaluate our short-term
137 (one-week-ahead) forecasts of reported mpox cases generated from EpiNow2: 1. one month into the
138 outbreak prior to exponential growth (June 13 and June 27); 2. during exponential growth (July 5); 3. near
139 the outbreak peak (July 27); and 4. during the declining phase (September 6, September 19, October 11,
140 and December 5). Ideally, the same day of the week would be used, but some historical versions of the
141 dataset were not available for this analysis and some dates fell on national holidays which may have
142 introduced additional delays. Like the real-time analyses, we used rash onset date as the first reference
143 date to define the reporting delay distribution for all eight time points, while the second reference date
144 changed over time (Table S1).

145 We used three metrics to evaluate the forecasts. Our primary metric was the weighted interval
146 score (WIS). For each of the eight time points considered, WIS was computed for each daily prediction
147 and averaged across the seven-day forecast. The WIS measures the consistency of a group of prediction
148 intervals with an observed value (probabilistic accuracy). The WIS is positive, and lower values

149 correspond to smaller error (Supplementary methods) (15). To evaluate the error in the forecast's point
150 estimate, we used the mean absolute error (MAE), which was computed as $MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$,
151 where y_i is the observed number of mpox cases on day i , \hat{y}_i is the median forecast on day i , and $N = 7$
152 (15). We also used prediction interval (PI) coverage rates, which check the degree to which the model
153 provides calibrated predictions. Coverage rates are calculated by determining the proportion of times the
154 50% or 90% PIs included the observed value (for example, a well-calibrated forecast would have a 50%
155 PI coverage close to 0.50. Also see Supplementary methods) (15).

156 We compared the performance of EpiNow2 with a naïve Bayesian generalized linear model
157 (GLM, Supplementary methods). We calculated the relative WIS and relative MAE for EpiNow2 and the
158 GLM as $\theta_{EpiNow2, GLM} = \frac{\text{mean score of EpiNow2}}{\text{mean score of GLM}}$, where the mean score is the average of the models'
159 performance (WIS or MAE) over all eight dates evaluated. If $\theta_{EpiNow2, GLM}$ was less than 1, that indicated
160 the forecasts generated by EpiNow2 had less error than the GLM, whereas $\theta_{EpiNow2, GLM} > 1$ indicated
161 EpiNow2 performed worse.

162 For both EpiNow2 and the GLM, we removed recent cases (defined as cases reported in the last 3
163 – 5 days) from the time series to adjust for right truncation of the data for all four outbreak phases (Table
164 S1). Forecasts were evaluated using mpox data as of March 16, 2023. We used the most recent version of
165 event date as the basis for the comparison.

166 **Results**

167 **Challenges of nowcasting/forecasting**

168 It is voluntary for jurisdictions to report mpox cases to CDC, with only minimal data needed to
169 submit a case report form (e.g., case ID and reporting jurisdiction) in part, because not all cases may be
170 reached or fully investigated. CDC asks jurisdictions to collect and report additional data variables to
171 achieve situational awareness and surveillance goals. The number of variables on the case report form
172 was decreased to reduce reporting burden. Despite these efforts, jurisdictional case surveillance systems
173 may not have aligned to CDC's requested data variables, and jurisdictions may choose to limit what data

174 are shared with CDC based on local reporting practices. Received data were subjected to additional
175 manual data cleaning to standardize formats and correct obvious data entry errors. As a result, even key
176 data variables such as demographic characteristics (e.g., age, race/ethnicity, HIV status, vaccination) were
177 not consistently available across all jurisdictions and time periods, precluding detailed sub-analyses. For
178 example, out of 29,921 cases in DCIPHER through December 31, 2022, 21,480 (71.8%) were missing
179 HIV status, 16,474 (55.1%) were missing smallpox vaccination, 3,913 (13.1%) were missing gender
180 identity, 3,172 (10.6%) were missing race, 2,928 (9.8%) were missing ethnicity, and 250 (0.8%) were
181 missing age. The timing and frequency of data submission varied between jurisdictions and changed over
182 the course of the outbreak. Some jurisdictions reported case data in near real-time whereas others
183 submitted a large number of cases all at once, the latter of which caused large, artificial spikes in the time-
184 series. There were instances of duplicate cases being reported from several jurisdictions which may be
185 attributed to the changes in reporting processes. Spurious cases at the end of the time series had to be
186 investigated (and usually removed) because they artificially inflated the nowcasts/forecasts. These data
187 issues required us to monitor the model output closely and modify the code as needed.

188 In early July 2022, reporting of mpox cases to CDC started to lag in some jurisdictions, especially
189 those most affected by the outbreak. These few jurisdictions were publicly reporting more cases on their
190 websites than what CDC had received reports for. This led to a lengthy case reconciliation process during
191 which case data uncertainty prevented it from being used for nowcasting/forecasting at the national level.
192 Also in July, an increasing proportion of cases were reported with missing symptom onset dates (from
193 26% to 53% for rash onset date between June 13 and July 5). To ensure each case had a date associated
194 with it for plotting epidemic curves, a new event date field was created which we started using for
195 nowcasting/forecasting. The new calculated date field selected the best available date among possible date
196 fields based on the following priority, ordered from most to least preferable: orthopoxvirus test date, date
197 of call to the call center, date the short CRF was created, and the long CRF timestamp (Figure S1). The
198 date in which the record was created was least preferred due to artificial spikes in the time series caused
199 by bulk uploading data. In September 2022, a different date was adopted by the response for reporting

200 case data. This date field was defined as the earliest among all available dates for a case including
201 symptom onset, which facilitated improved visualization of epidemic curves. However, this new date
202 presented challenges for its use in the EpiNow2 framework because the delay from symptom onset to this
203 new date would have more variation than the delay using the original event date, including a delay of zero
204 for some cases. Thus, we worked to create a new event date field specifically for nowcasting/forecasting
205 which was similar to the original event date. The definition was expanded to include dates available in
206 NNDSS data.

207 **Successes of nowcasting/forecasting**

208 During the case reconciliation process in July, publicly available data through health department
209 websites was used for subnational analyses [e.g., California (28) and New York City (29)]. We used
210 WebPlotDigitizer (30) to extract time series data from pdfs when the underlying data were not available
211 for download.

212 In October, we updated estimates of the serial interval for rash onset of 7.0 days (95% CrI 5.8 –
213 8.4) from 40 case pairs and incubation period for rash onset of 7.5 days (95% CrI 6.0 – 9.8) from 35 U.S.
214 case-patients and used those as model inputs for EpiNow2 (31). These data were obtained through the
215 collaboration of several U.S. jurisdictions on a special study. The estimated serial interval for the 2022
216 outbreak was on the lower end of the historical range observed in the Democratic Republic of Congo (7 –
217 23 days) (24).

218 **Adapting model output and communicating nowcasts/forecasts**

219 We adapted the presentation of our results for a scientific/technical audience and the general
220 public. The default plots from EpiNow2 included three panels: cases by date of report, cases by date of
221 infection, and R_t . Green represented estimates based on complete data, orange represented estimates based
222 on partial data, and purple represented the forecast (the default is seven days). Gray bars in the top panel
223 showed the actual time series of reported cases, while gray bars in the middle panel showed the back-
224 calculated infection time series. For the Technical Reports, Situational Reports, and response updates
225 meetings, we removed the middle panel (Figure 1) (26). For the website, we only showed R_t , removed the

226 forecast, and removed the 20% credible intervals to minimize confusion (Figure 2) (27). We included a
227 simple description of the plot that could be understood by non-experts. In accordance with CDC's Data
228 Modernization Initiative, a national effort aimed at modernizing state and national core data and
229 surveillance infrastructure (32), data for the underlying plots were made available for download as
230 comma-separated values (csv) files with the Technical Reports.

231 Sub-national analyses revealed some differences between regions regarding the start of the
232 outbreak, when it peaked, and how long it lasted (Figures 3 – 4). For example, Figure 4 demonstrates a
233 later introduction date and slightly longer tail for Texas compared to other jurisdictions.

234 **Performance assessment**

235 The GLM had lower WIS compared to EpiNow2 for early phase of the outbreak (Table 1);
236 however, during all other phases, EpiNow2 had a slight advantage. The relative WIS was 0.89 over all
237 eight time points considered, indicating that EpiNow2 had on average 11% less probabilistic error than
238 the GLM.

239 EpiNow2 had lower MAE than the GLM for six out of eight time points, but performance was
240 similar: the relative MAE was 0.96. In other words, EpiNow2 had only 4% less point error than the GLM.

241 Overall, predictions were moderately well calibrated. For the 90% PI, EpiNow2 achieved
242 coverage rates within 10% of the desired coverage level for seven out of eight time points compared to six
243 out of eight for the GLM. For the 50% PI, EpiNow2 achieved coverage rates within 10% of the desired
244 coverage level for only two out of eight time points compared to five out of eight for the GLM.

245 Qualitative results of the nowcasts/forecasts are shown in Figure S2. The 90% CrIs from
246 EpiNow2 were very wide for the early phase of the outbreak, while the GLM had large uncertainty
247 around the outbreak's peak. Both models underestimated reported mpox cases for the seventh time point
248 on October 11 which could be due to a discrepancy in the data available at the time versus the ground-
249 truth data (Figure S3). This time point had the lowest PI coverage rates.

250 **Discussion**

251 We performed nowcasting/forecasting to inform the U.S. response to the 2022 mpox outbreak in
252 real-time. Validation showed that the method implemented in EpiNow2 predicted case counts reasonably
253 well, but improvements are needed around key time periods such as the outbreak's peak. One reason that
254 the nowcasts/forecasts did not always align with reality is because the definition of event date changed
255 over time, while the study data were constructed the most recent version of the event date field. We found
256 a higher WIS for EpiNow2 in the early phase of the outbreak compared to the GLM which could be due
257 to choices of priors for parameters (e.g., wide intervals for R_t).

258 Subnational analyses allowed us to better understand the spatial heterogeneity of the epidemic
259 which may be attributed to differences between jurisdictions in terms of composition (e.g., population age
260 structure, density, and contact patterns) and public health activities (e.g., vaccination, surveillance
261 methods, frequency of testing) (33) as well as the timing and frequency of case reporting. One limitation
262 of nowcasting/forecasting at the subregional or jurisdictional level is that the effects of bulk uploads are
263 more apparent, resulting in greater uncertainty (wider credible intervals). Another limitation is that
264 movement between jurisdictions could have a greater impact on subnational estimates, as mobility is not
265 accounted for in our approach. Finally, some jurisdictions stopped reporting rash onset date, which
266 decreased the sample size available for estimating the reporting delay distribution over time.

267 *Data Quality*

268 Nowcasting/forecasting methods perform best when the underlying surveillance data are accurate,
269 timely, and complete, but they are often sub-optimal and variable as the outbreak evolves; while data may
270 improve as an outbreak progresses, they may re-deteriorate once the outbreak slows and intensity of effort
271 is low. Fortunately, the quality and frequency of data improved over the course of the U.S. mpox
272 outbreak. Communicating with specific jurisdictions about our priority dates for modeling improved data
273 quality. These prompts to the jurisdictions need to be continued regularly throughout the outbreak. Close
274 collaboration between epidemiologists/modelers and informaticians, including the use of an issue tracking
275 system in DCIPHER, also facilitated quick investigation and resolution of data errors.

276 *EpiNow2 Limitations*

277 The main limitation of EpiNow2 is its steep learning curve due to limited documentation of
278 package functions and few reports of its application to other outbreaks. Increasing commenting in the
279 code, creating more tutorials or vignettes, and developing a graphical user interface could help.

280 Another limitation is the long computing time required for the analyses. We were able to increase
281 computational efficiency by running the model on multiple cores in parallel, but the processing time
282 became particularly cumbersome if an analysis needed to be repeated. In the future, cloud-based
283 computing could be used to obtain more consistent and faster model run times.

284 There were also instances of unusually long run times whereby the first two Markov chain Monte
285 Carlo (MCMC) chains performed as expected, but subsequent chains never finished processing. Some
286 MCMC convergence issues were resolved by reducing the parameter fitting period (e.g., truncating the
287 beginning of the time series). One study reported that EpiNow2 estimates are more reliable when case
288 numbers at each time step are large and there are at least 14 timepoints without zeroes (34). Large daily
289 fluctuations and limited case counts could substantially affect model estimates, which should be
290 interpreted with caution.

291 Another limitation is that the method we used does not account for under-ascertainment, which
292 occurs when not all infections are diagnosed and reported as cases of the disease to the surveillance
293 system. The under-ascertainment rate is needed to understand the true burden of disease caused by the
294 outbreak; however, current estimates for the U.S. mpox outbreak are lacking. Indirect evidence from a
295 recent modeling study (35) suggests that 65% of mpox infections were diagnosed and reported in
296 Washington, D.C. However, the model was not designed to measure the under-ascertainment rate (Patrick
297 Clay, personal communication, March 10, 2023), and this quantity should be assessed by other methods
298 (e.g., models specifically designed to assess under-ascertainment, serological surveys, and community-
299 based surveys).

300 ***Strategies for Forecasting the Next Outbreak***

301 For the next outbreak, it is important for CDC to develop strategies for regularly capturing and
302 storing snapshots of surveillance data which remain easily accessible for systematic analysis. For routine

303 case-based surveillance of notifiable diseases, such as rabies, most analyses are performed only after the
304 data have undergone a rigorous and routine reconciliation and closeout process by data submitters with
305 further validation by CDC surveillance epidemiologists; however, timely outbreak response decision
306 support does not allow for such processes. Instead, jurisdictions are asked to submit available case data in
307 near real-time and submit additional data or corrections to data entry errors as time and resources permit.
308 Snapshots of the surveillance data were saved in an ad hoc manner (by exporting data on a particular day
309 and saving a csv file locally), and as a consequence, a complete history of the data is not available,
310 especially around key points in the outbreak, such as the peak. A complete history would help to
311 understand key delay distributions and other quirks (e.g., backfilling and revision of reference dates)
312 involved in the data-generating process. Understanding the data generating process is crucial for the
313 improvement of methods and tools for nowcasting/forecasting and aligns with one of the five priorities of
314 CDC's Data Modernization Initiative (Accelerating Data for Action: Tapping into more data sources,
315 promoting health equity, and increasing capacities for scalable outbreak response, forecasting, and
316 predictive analytics) (32). In the future, the process of saving snapshots of the data could be automated.

317 Ensemble models have been used for a variety of infectious disease outbreaks, such as COVID-
318 19 (15, 36), Zika (37), influenza (38), and Ebola (39). Ensembles combine predictions from several
319 models that use different methodology and sometimes input data. Because some models overpredict,
320 while others underpredict, ensemble models often outperform individual models over time. In the future,
321 we may consider using at least two simpler models and comparing them.

322 One potentially useful addition to EpiNow2 and other currently available tools for
323 nowcasting/forecasting outbreaks would be flexibility in handling dates. We frequently encountered
324 missing dates for cases in the mpox surveillance data. Ideally, a method or tool would be able to keep
325 track of multiple dates for a case and estimate missing dates based on the full distribution of dates across
326 all cases. EpiNowcast is a new hierarchical nowcasting package that enables more flexibility in adjusting
327 for truncated data (40). Novel nowcasting approaches use hierarchical generalized additive models, which
328 can provide even more flexibility to modify the model in real-time to the evolving data environment (41).

329 Another improvement would be to reduce the time required to run the analyses. Rather than focusing on
330 the efficiency of the MCMC algorithm, computation time could be reduced if the model only needed to be
331 run on the new data. Finally, the imputed time series of cases by symptom onset date would be a useful
332 data visualization output that is not currently available in EpiNow2. As described above, defining the date
333 field for the presentation of epidemic curves was a challenge in the mpox outbreak and having an imputed
334 symptom onset date for each case would have been useful for comparison purposes.

335 CFA played an important advisory role in our nowcasting/forecasting efforts. CFA produces
336 models and forecasts to characterize the state of an outbreak and its course, inform public health decision
337 makers on potential consequences of deploying control measures, and support innovation to continuously
338 improve the science of outbreak analytics and modeling (42). In the future, CFA plans to create new tools
339 for outbreak analysis and modeling. CFA could also serve as a link between CDC modelers and
340 jurisdictions with modeling capacity to share experiences and code. Technical Reports represent a new
341 way for CDC to share timely information with the federal government, state and local leaders, and
342 scientists in academia and industry. Technical Reports have been well received within and outside CDC
343 (43-45) and their publication aligns with CDC's current restructuring efforts aimed at making the agency
344 more response ready, including sharing science and data faster (46).

345 **Conclusion**

346 Real-time estimation of R_t as well as nowcasting/forecasting is one method for determining the
347 extent to which current public health measures are effective and/or need to be modified but is subject to
348 limitations. The quality and timeliness of reported data pose challenges to these analyses. Ease of use,
349 model computing time, and ability to handle multiple dates are priorities for consideration in the
350 development of future nowcasting/forecasting tools. A naïve model may be superior to a complex one,
351 such as EpiNow2, during the early phase of an outbreak when data scarcity causes R_t to be largely
352 unconstrained, especially once reporting delays are considered. Future outbreak response activities could
353 be enhanced through inclusion of clear and consistent communication about modeling outputs as well as
354 close collaboration between modeling and informatics/data teams.

356 **Tables**

357 **Table 1.** Evaluation of short-term (one-week-ahead) forecasts of reported mpox cases during the 2022 U.S. outbreak. For WIS, bold indicates
 358 where one model performed better than the other.

Forecast date	Outbreak phase	EpiNow2				Bayesian GLM with negative binomial			
		90% PI coverage	50% PI coverage	WIS	MAE	90% PI coverage	50% PI coverage	WIS	MAE
Monday, June 13	Early	1	0.71	33.8	4.3	0.86	0.57	6.4	5.1
Monday, June 27	Early	0.86	0	34.0	29.0	0.86	0.14	24.7	26.9
Tuesday, July 5	Exponential growth	0.86	0.71	35.6	32.4	1	0.57	38.5	34.1
Wednesday, July 27	Peak	1	0.57	137.0	107.9	1	0.57	184.3	94.4
Tuesday, September 6	Decline	0.86	0.71	107.6	84.3	0.86	0.43	117.1	86.8
Monday, September 19	Decline	0.86	0.43	61.2	52.4	0.71	0.43	81.8	65.9
Tuesday, October 11	Decline	0.43	0.29	30.6	33.4	0.57	0.14	38.3	41.9
Monday, December 5	Decline	1	0.71	4.3	3.1	1	0.71	6.4	4.6

359 PI: prediction interval; WIS: weighted interval score; MAE: mean absolute error; GLM: generalized linear model

360 **Figures**

361 **Figure 1.** Effective reproduction number estimates for the U.S. 2022 mpox outbreak intended for a
362 technical/scientific audience. The top panel shows estimates of cases by date of report with actual cases
363 shown by gray bars. The bottom panel shows estimates of the effective reproduction number by date. In
364 all panels, shaded regions reflect 90%, 50%, and 20% credible intervals in order from lightest to darkest.
365 Green shows estimates, red shows estimates based on partial data, and purple shows forecasts. Event date
366 is determined by a hierarchy across the different data streams where priority is given to diagnosis date,
367 orthopoxvirus test date, orthopoxvirus test confirmation date, case investigation start date, orthopoxvirus
368 sample collection date, date of call to CDC call center, report date (to public health department, county, or
369 state), date CDC announced case, and the date the case was entered into DCIPHER, in that order.

370
371 **Figure 2.** Effective reproduction number (R_t) estimates for the U.S. 2022 mpox outbreak intended for the
372 general public. The graph shows the R_t estimation over time based on complete data (gray) or partial data
373 (blue). The most recent data are considered incomplete due to delays in reporting mpox cases. As a result,
374 there is more uncertainty associated with the most recent R_t estimates. $R_t > 1$ means the epidemic is
375 growing. $R_t < 1$ means the epidemic is shrinking. Shading represents the 50% and 90% credible intervals
376 (uncertainty in the estimates)

377
378 **Figure 3.** Effective reproduction number estimates for the 2022 mpox outbreak in four U.S. Census
379 regions. The left panels show estimates of cases by date of report with actual cases shown by gray bars.
380 The right panels show estimates of the effective reproduction number by date. In all panels, shaded
381 regions reflect 90%, 50%, and 20% credible intervals in order from lightest to darkest. Green shows
382 estimates, red shows estimates based on partial data, and purple shows forecasts. Event date is determined
383 by a hierarchy across the different data streams where priority is given to diagnosis date, orthopoxvirus
384 test date, orthopoxvirus test confirmation date, case investigation start date, orthopoxvirus sample
385 collection date, date of call to CDC call center, report date (to public health department, county, or state),
386 date CDC announced case, and the date the case was entered into DCIPHER, in that order.

387
388 **Figure 4.** Effective reproduction number estimates of the 2022 mpox outbreak for the six jurisdictions in
389 the U.S. with the highest case counts. The left panels show estimates of cases by date of report with actual
390 cases shown by gray bars. The right panels show estimates of the effective reproduction number by date.
391 In all panels, shaded regions reflect 90%, 50%, and 20% credible intervals in order from lightest to
392 darkest. Green shows estimates, red shows estimates based on partial data, and purple shows forecasts.
393 Event date is determined by a hierarchy across the different data streams where priority is given to
394 diagnosis date, orthopoxvirus test date, orthopoxvirus test confirmation date, case investigation start date,
395 orthopoxvirus sample collection date, date of call to CDC call center, report date (to public health
396 department, county, or state), date CDC announced case, and the date the case was entered into
397 DCIPHER, in that order.

398 **Data availability**

399 Data and code to run the nowcasts/forecasts and perform model validation will be available on GitHub
400 following publication in a peer-reviewed journal.

401

402 **Acknowledgements**

403 We thank all public health professionals involved in reporting mpox cases to CDC. We also acknowledge
404 the Mpox Response Data Analytics and Visualization Task Force Informatics Team for data management
405 and support, and we thank Dr. Sam Abbott for helpful discussions about EpiNow2 methods.

406

407 **Declaration of interest**

408 The authors declare the following financial interests/personal relationships which may be considered as
409 potential competing interests: Kelly Charniga reports a relationship with Systems Planning and Analysis
410 Inc that includes: consulting or advisory.

411

412 **Disclaimer**

413 The findings and conclusions in this report are those of the authors and do not necessarily represent the
414 official position of the Centers for Disease Control and Prevention, U.S. Department of Health and
415 Human Services.

416

417 **Ethics statement**

418 This activity was reviewed by CDC and was conducted consistent with applicable federal law and CDC
419 policy (45 C.F.R. part 46, 21 C.F.R. part 56; 42 U.S.C. Sect. 241(d); 5 U.S.C. Sect. 552a; 44 U.S.C. Sect.
420 3501 et seq).

421

422 **Funding**

423 No specific funding was obtained for this work.

424

425 **Author contributions**

426 Conceptualization (KC, YN), data curation (KC, JA), formal analysis (KC, ZJM), investigation (KC,
427 ZJM), methodology (KC, ZJM, JA), project administration (KC), software (KC, ZJM, NBM), supervision
428 (JA, YN, IHS), validation (KC, ZJM), visualization (KC, ZJM), writing – original draft preparation (KC,
429 ZJM), writing – review & editing (KC, ZJM, NBM, JA, YN, IHS).

430

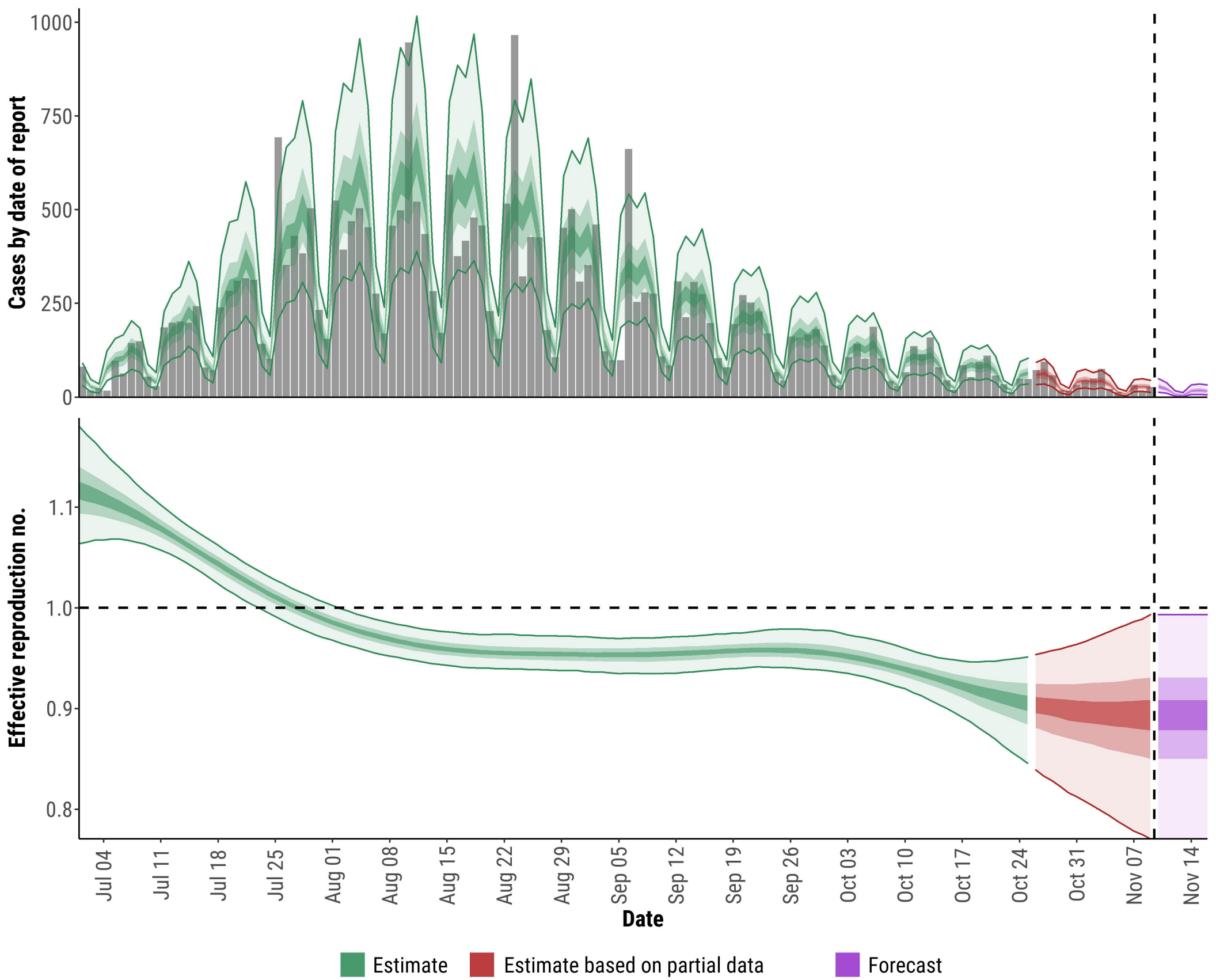
431 **References**

- 432 1. World Health Organization. WHO Director-General declares the ongoing monkeypox outbreak a
433 Public Health Emergency of International Concern 2022 [updated July 23, 2022. Available from:
434 [https://www.who.int/europe/news/item/23-07-2022-who-director-general-declares-the-ongoing-](https://www.who.int/europe/news/item/23-07-2022-who-director-general-declares-the-ongoing-monkeypox-outbreak-a-public-health-event-of-international-concern)
435 [monkeypox-outbreak-a-public-health-event-of-international-concern](https://www.who.int/europe/news/item/23-07-2022-who-director-general-declares-the-ongoing-monkeypox-outbreak-a-public-health-event-of-international-concern).
- 436 2. U.S. Centers for Disease Control and Prevention. 2022 Monkeypox Outbreak Global Map 2022
437 [updated November 15, 2022. Available from:
438 <https://www.cdc.gov/poxvirus/monkeypox/response/2022/world-map.html>.
- 439 3. Beer E, Rao V. A systematic review of the epidemiology of human monkeypox outbreaks and
440 implications for outbreak strategy. *PLoS Negl Trop Dis*. 2019;13(10):e0007791.
- 441 4. U.S. Centers for Disease Control and Prevention. Signs and Symptoms 2022 [updated October
442 18, 2022. Available from: <https://www.cdc.gov/poxvirus/monkeypox/symptoms/index.html>.
- 443 5. WHO/ECDC. Interim advice for public health authorities on summer events during the
444 monkeypox outbreak in Europe, 2022 2022 [updated June 14, 2022. Available from:
445 [https://www.ecdc.europa.eu/sites/default/files/documents/Interim-advice-for-public-health-](https://www.ecdc.europa.eu/sites/default/files/documents/Interim-advice-for-public-health-authorities-on-summer-events-mpx.pdf)
446 [authorities-on-summer-events-mpx.pdf](https://www.ecdc.europa.eu/sites/default/files/documents/Interim-advice-for-public-health-authorities-on-summer-events-mpx.pdf).
- 447 6. UK Health Security Agency. Investigation into monkeypox outbreak in England: technical briefing
448 1 2022 [updated September 23, 2022. Available from:
449 [https://www.gov.uk/government/publications/monkeypox-outbreak-technical-briefings/investigation-](https://www.gov.uk/government/publications/monkeypox-outbreak-technical-briefings/investigation-into-monkeypox-outbreak-in-england-technical-briefing-1#:~:text=Without%20correcting%20for%20right%2Dtruncation,8.9)%2C%20see%20Figure%202a)
450 [into-monkeypox-outbreak-in-england-technical-briefing-](https://www.gov.uk/government/publications/monkeypox-outbreak-technical-briefings/investigation-into-monkeypox-outbreak-in-england-technical-briefing-1#:~:text=Without%20correcting%20for%20right%2Dtruncation,8.9)%2C%20see%20Figure%202a)
451 [1#:~:text=Without%20correcting%20for%20right%2Dtruncation,8.9\)%2C%20see%20Figure%202a](https://www.gov.uk/government/publications/monkeypox-outbreak-technical-briefings/investigation-into-monkeypox-outbreak-in-england-technical-briefing-1#:~:text=Without%20correcting%20for%20right%2Dtruncation,8.9)%2C%20see%20Figure%202a).
- 452 7. Thompson J, McClure R, Scott N, Hellard M, Abey Suriya R, Vidanaarachchi R, et al. A framework
453 for considering the utility of models when facing tough decisions in public health: a guideline for policy-
454 makers. *Health Research Policy and Systems*. 2022;20:107.
- 455 8. Buchwald A, Adams J, Bortz D, Carlton E. Infectious Disease Transmission Models to Predict,
456 Evaluate, and Improve Understanding of COVID-19 Trajectory and Interventions. *Annals of the American*
457 *Thoracic Society*. 2020;17(10):1204-6.
- 458 9. Madewell Z, Pastore Y Piontti A, Zhang Q, Burton N, Yang Y, Longini I, et al. Using simulated
459 infectious disease outbreaks to inform site selection and sample size for individually randomized vaccine
460 trials during an ongoing epidemic. *Clinical Trials*. 2021;18(5):630-8.

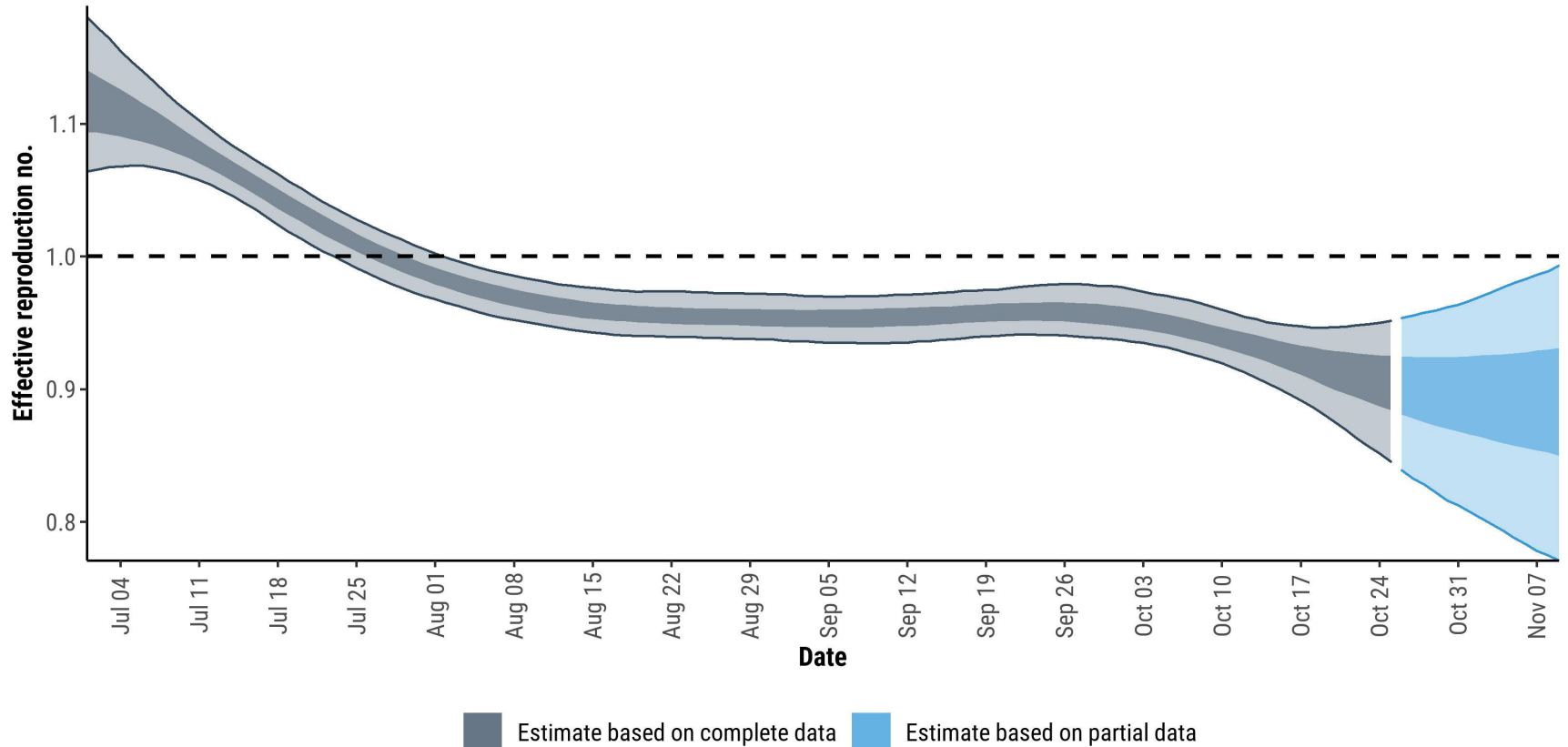
- 461 10. Lloyd-Smith J, Funk S, McLean A, Riley S, Wood J. Nine challenges in modelling the emergence of
462 novel pathogens. *Epidemics*. 2015;10:35-9.
- 463 11. Madewell Z, Dean N, Berlin J, Coplan P, Davis K, Struchiner C, et al. Challenges of evaluating and
464 modelling vaccination in emerging infectious diseases. *Epidemics*. 2021;37:100506.
- 465 12. Chakraborty T, Ghosh I, Mahajan T, Arora T. Nowcasting of COVID-19 Confirmed Cases:
466 Foundations, Trends, and Challenges. In: Azar A, Hassanien A, editors. *Modeling, Control and Drug
467 Development for COVID-19 Outbreak Prevention*. Studies in Systems, Decision and Control: Springer,
468 Cham; 2022.
- 469 13. Abbott S, Hellewell J, Thompson R, Sherratt K, Gibbs H, Bosse N, et al. Estimating the time-
470 varying reproduction number of SARS-CoV-2 using national and subnational case counts [version 2; peer
471 review: 1 approved, 1 approved with reservations]. *Wellcome Open Research*. 2020;5:112.
- 472 14. Cori A, Ferguson N, Fraser C, Cauchemez S. A New Framework and Software to Estimate Time-
473 Varying Reproduction Numbers During Epidemics. *American Journal of Epidemiology*. 2013;178(9):1505-
474 12.
- 475 15. Cramer E, Ray E, Lopez V, Bracher J, Brennen A, Rivadeneira A, et al. Evaluation of individual and
476 ensemble probabilistic forecasts of COVID-19 mortality in the United States. *PNAS*.
477 2022;119(15):e2113561119.
- 478 16. Arslan S, Ozdemir M, Ucar A. Nowcasting and Forecasting the Spread of COVID-19 and
479 Healthcare Demand in Turkey, a Modeling Study. *Front Public Health*. 2021;8:575145.
- 480 17. Birrell P, Blake J, van Leeuwen E, Gent N, De Angelis D. Real-time nowcasting and forecasting of
481 COVID-19 dynamics in England: the first wave. *Philos Trans R Soc Lond B Biol Sci*.
482 2021;376(1829):20200279.
- 483 18. Spreco A, Jöud A, Eriksson O, Soltesz K, Källström R, Dahlström Ö, et al. Nowcasting (Short-Term
484 Forecasting) of COVID-19 Hospitalizations Using Syndromic Healthcare Data, Sweden, 2020. *Emerg Infect
485 Dis*. 2022;28:564-71.
- 486 19. Biggerstaff M, Slayton R, Johansson M, Butler J. Improving Pandemic Response: Employing
487 Mathematical Modeling to Confront Coronavirus Disease 2019. *Clinical Infectious Diseases*.
488 2021;74(5):913-7.
- 489 20. U.S. Department of Health and Human Services. Privacy Impact Assessment 2020 [updated
490 March 26, 2020. Available from: <https://www.hhs.gov/sites/default/files/cdc-dciphcr-whr.pdf>.
- 491 21. Cori A, Cauchemez S, Ferguson N, Fraser C, Dahlgvist E, Demarsh P, et al. EpiEstim: Estimate
492 Time Varying Reproduction Numbers from Epidemic Curves. 2.2-4 ed. CRAN2021.
- 493 22. Jombart T, Cori A, Nouvellet P, Skaup J, Kamvar Z, Taylor T. earlyR: Estimation of Transmissibility
494 in the Early Stages of a Disease Outbreak. 0.0.5 ed: CRAN; 2020.
- 495 23. Abbott S, Hellewell J, Sherratt K, Gostic K, Hickson J, Badr H, et al. EpiNow2: Estimate real-time
496 case counts and time-varying epidemiological parameters. 1.3.2 ed: EpiForecasts; 2021.
- 497 24. Fenner F, Henderson D, Arita I, Jezek Z, Ladnyi I. *Smallpox and its Eradication*1988. Available
498 from: <https://apps.who.int/iris/handle/10665/39485>.
- 499 25. Nash R, Nouvellet P, Cori A. Real-time estimation of the epidemic reproduction number: Scoping
500 review of the applications and challenges. *PLOS Digit Health*. 2022;1(6):e0000052.
- 501 26. U.S. Centers for Disease Control and Prevention. Monkeypox Technical Reports 2022 [updated
502 October 27, 2022. Available from: [https://www.cdc.gov/poxvirus/monkeypox/cases-data/technical-
503 report.html](https://www.cdc.gov/poxvirus/monkeypox/cases-data/technical-report.html).
- 504 27. U.S. Centers for Disease Control and Prevention. Outbreak Reproduction Number Estimates
505 2022 [updated November 14, 2022. Available from: [https://www.cdc.gov/poxvirus/monkeypox/cases-
506 data/case-nowcasting.html](https://www.cdc.gov/poxvirus/monkeypox/cases-data/case-nowcasting.html).
- 507 28. California Department of Public Health. Monkeypox Data in California 2022 [updated November
508 10, 2022. Available from: <https://www.cdph.ca.gov/Programs/CID/DCDC/Pages/Monkeypox-Data.aspx>.

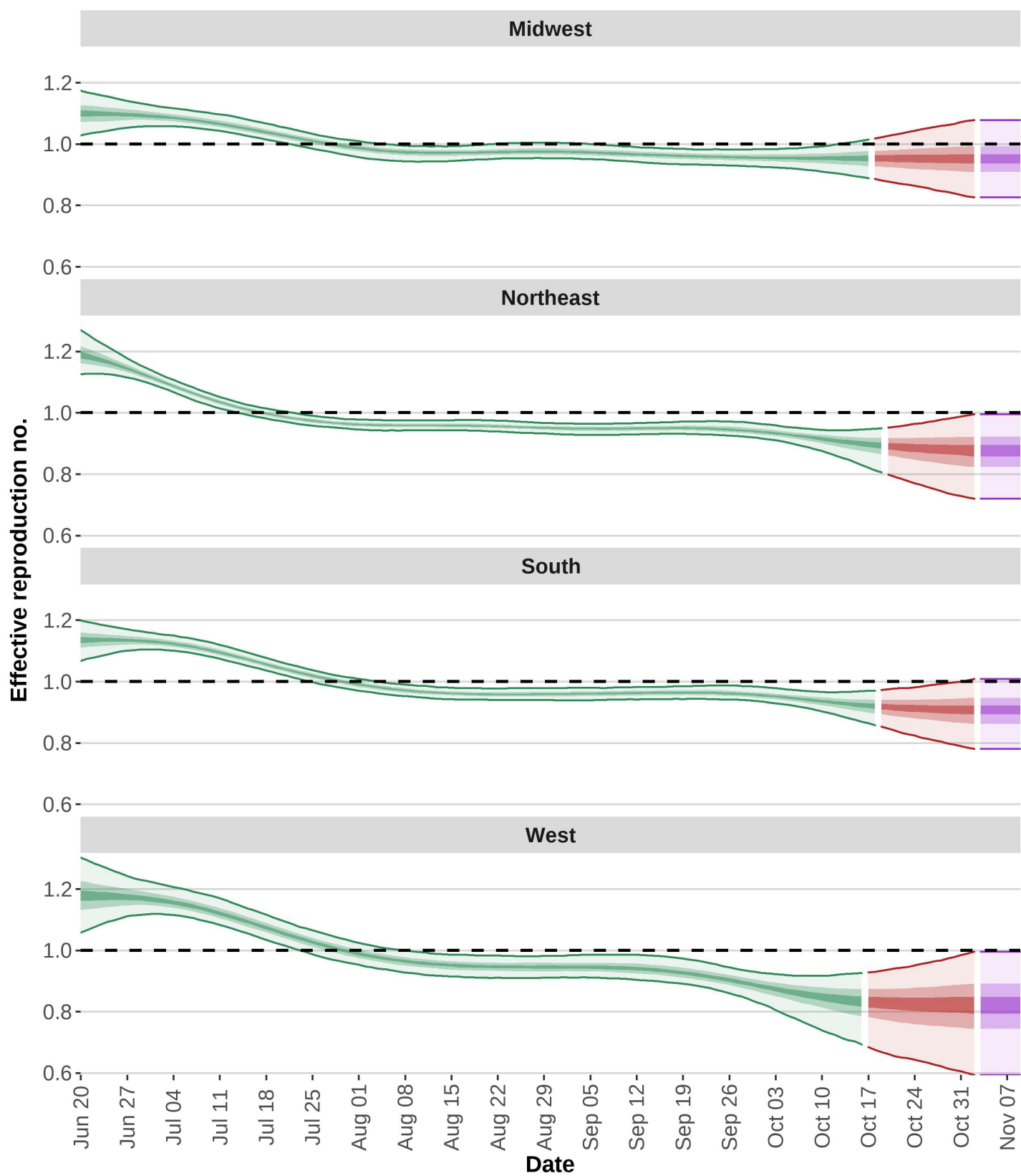
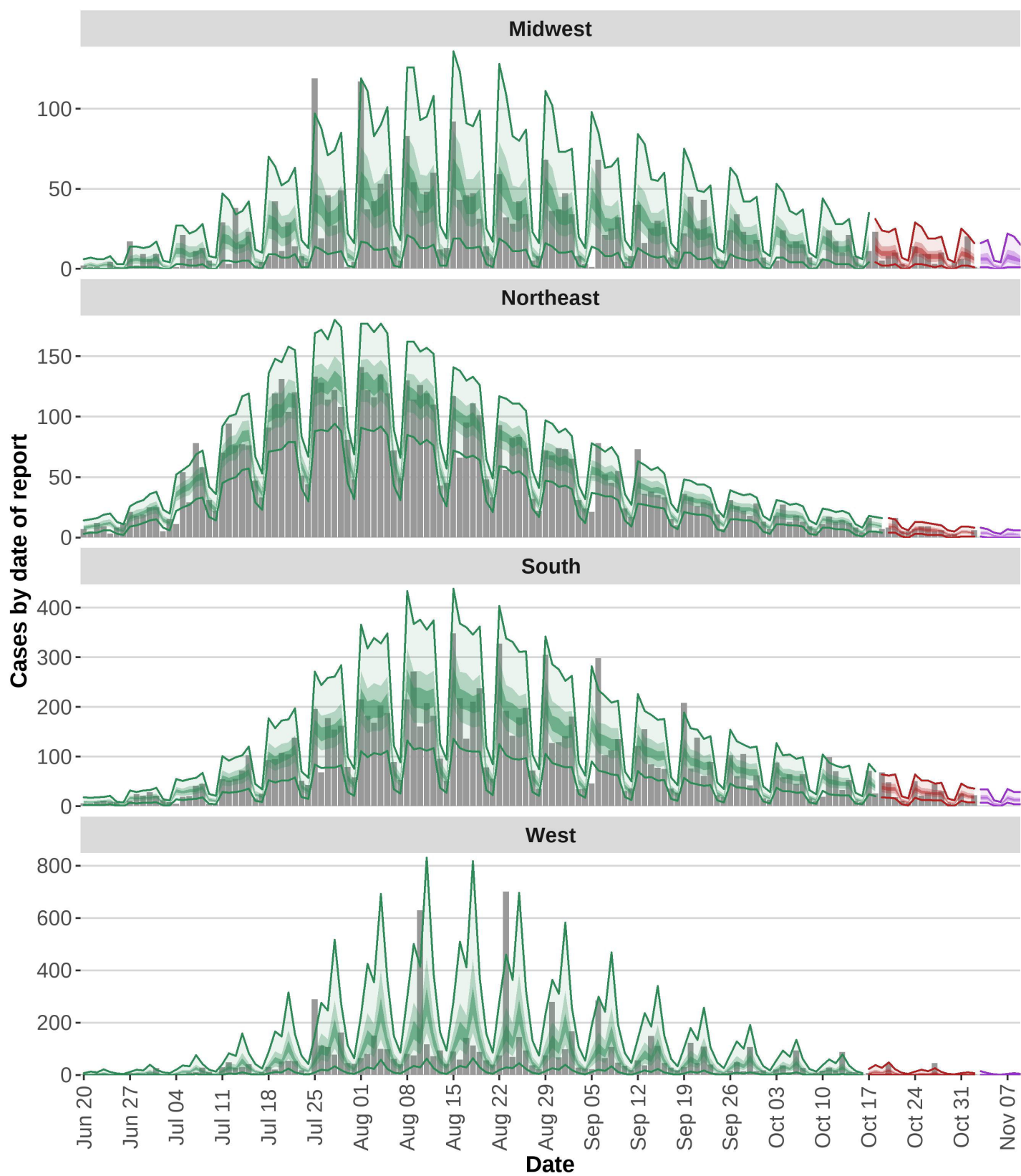
- 509 29. New York City Department of Health and Mental Hygiene. 2022 Monkeypox Surveillance Data
510 2022 [updated July 19, 2022. Available from:
511 [https://www1.nyc.gov/assets/doh/downloads/pdf/monkeypox/data-daily-cases-demographics-](https://www1.nyc.gov/assets/doh/downloads/pdf/monkeypox/data-daily-cases-demographics-072122.pdf)
512 [072122.pdf](https://www1.nyc.gov/assets/doh/downloads/pdf/monkeypox/data-daily-cases-demographics-072122.pdf).
- 513 30. Rohatgi A. WebPlotDigitizer. Vers. 4.5 ed. Pacifica, California, USA2021.
- 514 31. Madewell Z, Charniga K, Masters N, Asher J, Fahrenwald L, Still W, et al. Serial interval and
515 incubation period estimates of monkeypox virus infection in 12 U.S. jurisdictions, May – August 2022.
516 *Emerg Infect Dis.* 2023;29(4):818-21.
- 517 32. U.S. Centers for Disease Control and Prevention. What is the Data Modernization Initiative?
518 2022 [updated October 11, 2022. Available from: [https://www.cdc.gov/surveillance/projects/dmi-](https://www.cdc.gov/surveillance/projects/dmi-initiative/index.html#:~:text=CDC%E2%80%99s%20Data%20Modernization%20Initiative%20%28DMI%29%20is%20at%20the,data%20reporting%2C%20management%2C%20and%20analytics%20across%20public%20health)
519 [initiative/index.html#:~:text=CDC%E2%80%99s%20Data%20Modernization%20Initiative%20%28DMI%29%20is%20at%20the,data%20reporting%2C%20management%2C%20and%20analytics%20across%20pu-](https://www.cdc.gov/surveillance/projects/dmi-initiative/index.html#:~:text=CDC%E2%80%99s%20Data%20Modernization%20Initiative%20%28DMI%29%20is%20at%20the,data%20reporting%2C%20management%2C%20and%20analytics%20across%20public%20health)
520 [blic%20health](https://www.cdc.gov/surveillance/projects/dmi-initiative/index.html#:~:text=CDC%E2%80%99s%20Data%20Modernization%20Initiative%20%28DMI%29%20is%20at%20the,data%20reporting%2C%20management%2C%20and%20analytics%20across%20public%20health).
- 521
- 522 33. De Nicola G, Schneble M, Kauermann G, Berger U. Regional now- and forecasting for data
523 reported with delay: toward surveillance of COVID-19 infections. *Adv Stat Anal.* 2022;106(3):407-26.
- 524 34. Binny R, Lustig A, Hendy S, Maclaren O, Ridings K, Vattiato G, et al. Real-time estimation of the
525 effective reproduction number of SARS-CoV-2 in Aotearoa New Zealand. *PeerJ.* 2022;10:e14119.
- 526 35. Clay P, Asher J, Carnes N, Copen C, Delaney K, Payne D, et al. Modelling the impact of
527 vaccination and sexual behavior change on reported cases of mpox in Washington D.C. *medRxiv.* 2023.
- 528 36. Dean N, Pastore Y Piontti A, Madewell Z, Cummings D, Hitchings M, Joshi K, et al. Ensemble
529 forecast modeling for the design of COVID-19 vaccine efficacy trials. *Vaccine.* 2020;38(46):7213-6.
- 530 37. Oidtman R, Omodei E, Kraemer M, Castañeda-Orjuela C, Cruz-Rivera E, Misnaza-Castrillón S, et
531 al. Trade-offs between individual and ensemble forecasts of an emerging infectious disease. *Nature*
532 *Communications.* 2021;12:5379.
- 533 38. Reich N, McGowan C, Yamana T, Tushar A, Ray E, Osthus D, et al. Accuracy of real-time multi-
534 model ensemble forecasts for seasonal influenza in the U.S. *PLOS Computational Biology.*
535 2019;15(11):e1007486.
- 536 39. Viboud C, Sun K, Gaffey R, Ajelli M, Fumanelli L, Merler S, et al. The RAPIDD ebola forecasting
537 challenge: Synthesis and lessons learnt. *Epidemics.* 2018;22:13-21.
- 538 40. Abbott S, Lison A, Funk S. epinowcast: Flexible hierarchical nowcasting. Zenodo; 2021.
- 539 41. Overton C, Abbott S, Christie R, Cumming F, Day J, Jones O, et al. Nowcasting the 2022 mpox
540 outbreak in England. *arXiv.* 2023.
- 541 42. U.S. Centers for Disease Control and Prevention. Center for Forecasting and Outbreak Analytics
542 2022 [updated October 28, 2022. Available from: [https://www.cdc.gov/forecast-outbreak-](https://www.cdc.gov/forecast-outbreak-analysis/index.html)
543 [analysis/index.html](https://www.cdc.gov/forecast-outbreak-analysis/index.html).
- 544 43. Daskalakis D. You know that you are a #publichealth nerd when you wake up on a gloomy
545 Saturday to read @CDCgov
- 546 Twitter; 2022 [updated October 1, 2022. [Tweet]]. Available from:
547 https://twitter.com/dr_demetre/status/1576196874079735808?s=48&t=9NVB0dQPIU9S4J58AJJfw.
- 548 44. Fox S. These monkeypox technical reports from the CDC are [three fire emojis]: Twitter; 2022
549 [updated November 8, 2022. [Tweet]]. Available from:
550 <https://twitter.com/FoxandtheFlu/status/1589978622026461184>.
- 551 45. Rivers C. Changing trends in monkeypox epidemiology: Substack; 2022 [updated September 2,
552 2022. Available from: [https://caitlinrivers.substack.com/p/changing-trends-in-monkeypox-](https://caitlinrivers.substack.com/p/changing-trends-in-monkeypox-epidemiology)
553 [epidemiology](https://caitlinrivers.substack.com/p/changing-trends-in-monkeypox-epidemiology).

554 46. U.S. Centers for Disease Control and Prevention. CDC Moving Forward Summary Report 2022
555 [updated September 1, 2022December 27, 2022]. Available from:
556 <https://www.cdc.gov/about/organization/cdc-moving-forward-summary-report.html>.
557

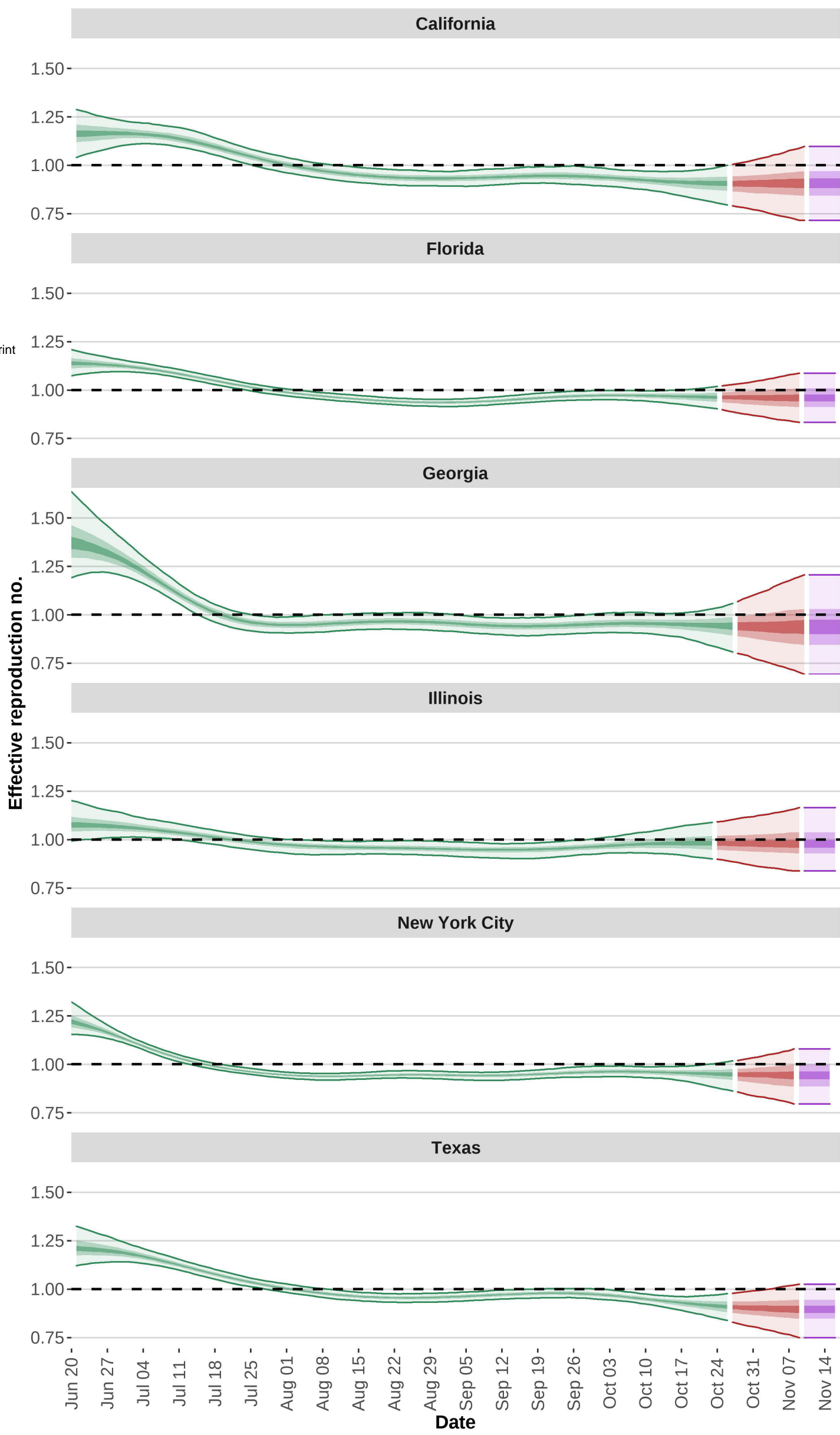
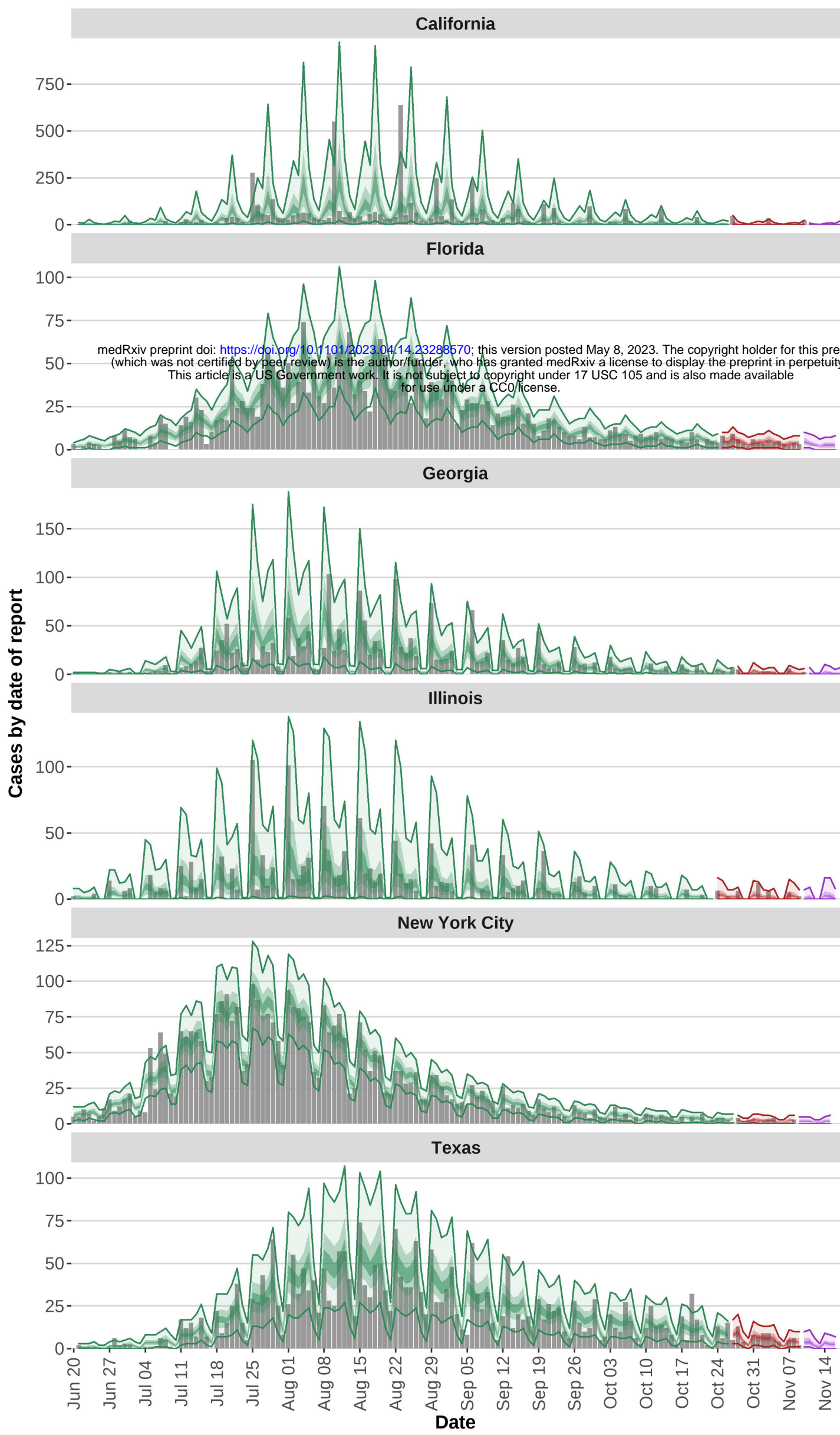


Average number of secondary infections transmitted from a primary infection





■ Estimate
 ■ Estimate based on partial data
 ■ Forecast



medRxiv preprint doi: <https://doi.org/10.1101/2023.04.14.23288570>; this version posted May 8, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

█ Estimate
 █ Estimate based on partial data
 █ Forecast