

1 The impact of local vaccine coverage and recent incidence
2 on measles transmission in France between 2009 and 2018

3 Authors:

4 Alexis Robert^{1,2*}, Adam J. Kucharski^{1,2}, Sebastian Funk^{1,2}

5 Affiliations:

6 1. Centre for the Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical
7 Medicine, Keppel Street, London, UK

8 2. Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine,
9 Keppel Street, London, UK

10

11 *Corresponding author

12 E-mail:alexis.robert@lshtm.ac.uk

13 Abstract

14 Background

15 Despite high levels of vaccine coverage, sub-national heterogeneity in immunity to measles can create
16 pockets of susceptibility, which are hard to detect and may result in long-lasting outbreaks. The
17 elimination status defined by the World Health Organization aims to identify countries where the virus
18 is no longer circulating and can be verified after 36 months of interrupted transmission. However, since
19 2018, numerous countries have lost their elimination status soon after reaching it, showing that the
20 indicators used to define elimination may not be predictive of lower risks of outbreaks.

21 Methods and Findings

22 We quantified the impact of local vaccine coverage and recent levels of incidence on the dynamics of
23 measles in each French department between 2009 and 2018, using mathematical models based on the
24 'Epidemic-Endemic' regression framework. High values of local vaccine coverage were associated with
25 fewer imported cases and lower risks of local transmissions. Regions that had recently reported high
26 levels of incidence were also at a lower risk of local transmission, potentially due to additional immunity
27 accumulated during these recent outbreaks. Therefore, all else being equal, the risk of local
28 transmission was not lower in areas fulfilling the elimination criteria (i.e., low recent incidence). After
29 fitting the models using daily case counts, we used the parameters' estimates to simulate the effect of
30 variations in the vaccine coverage and recent incidence on future transmission. A decrease of 3% in the
31 three-year average vaccine uptake led to a five-fold increase in the number of cases simulated in a year
32 on average.

33 Conclusions

34 Spatiotemporal variation in vaccine coverage because of disruption of routine immunisation
35 programmes, or lower trust in vaccines, can lead to large increases in both local and cross regional
36 transmission. The association found between local vaccine coverage and incidence suggests that,
37 although regional vaccine uptake can be hard to collect and unreliable because of population
38 movements, it can provide insights into the risks of imminent outbreak. Periods of low local measles
39 incidence were not indicative of a decrease in the risks of local transmission. Therefore, the incidence
40 indicator used to define the elimination status was not consistently associated with lower risks of
41 measles outbreak in France. More detailed models of local immunity levels or subnational
42 seroprevalence studies may yield better estimates of local risk of measles outbreaks.

43 Introduction

44 Immunity against infectious diseases accumulates following infection and, if a vaccine is available,
45 routine immunisation programs and vaccination campaigns. Measles is highly infectious and can cause
46 large outbreaks in populations with low immunity [1,2]. Therefore, high levels of vaccine coverage are
47 required to minimise the risks of outbreaks [3]. Furthermore, vaccine uptake must be homogeneously
48 high across the territory to avoid local transmission sustained by regional discrepancies [4,5]. The large-
49 scale implementation of routine immunisation programs led to a drastic reduction in measles cases
50 worldwide, and measles was targeted for elimination in five World Health Organization (WHO) Regions
51 by 2020 under the Global Vaccine Action Plan 2011-2020 [6].

52 Elimination status, as defined by the WHO, refers to “the absence of endemic measles transmission for
53 ≥ 12 months in the presence of a well-performing surveillance system” in a given country or region, and
54 is verified “after 36 months of interrupted endemic measles virus transmission” [7]. Although imported
55 cases, or cases directly related to importations could still be expected, there should be no continuous
56 transmission persisting over a long period of time in a region where measles was eliminated. A given
57 WHO region can declare measles eliminated when all countries in the region document interruption of
58 endemic transmission for more than 36 months.

59 Recently, several countries had their elimination status revoked following large outbreaks less than five
60 years after it was verified. For instance, the United Kingdom achieved elimination in 2017, and lost the
61 status in 2019 along with Albania, Czechia, Greece, Venezuela, and Brazil [8,9]. In these countries,
62 interruption of transmission during a few years was not indicative of reduced risks of major outbreaks.
63 Such occurrences can be explained by several factors, such as a replenishment of susceptible individuals
64 after years without transmission, or importations of cases into subnational areas with lower levels of
65 immunity caused by heterogeneity in vaccine coverage [10–13]. The number and geographical
66 distribution of the susceptible individuals is not routinely monitored in most countries given the
67 perceived cost and logistical challenges of large serological surveys, yet it is a main predictor of outbreak
68 risk [3]. Local values of vaccine coverage can be an alternative measure of heterogeneity, but they are
69 not always available and can be outdated because of the mobility between regions. Furthermore, they
70 only describe vaccine-induced immunity, and therefore ignore the immunity caused by previous
71 outbreaks. In this study, we aim to i) estimate the impact of recent local transmission and local vaccine
72 coverage on the current risk of outbreaks, and the changes in transmission dynamics that would result
73 from variations in these factors, and ii) identify the areas most at-risk for local transmission using France
74 as a case study.

75 To do so, we implemented an Epidemic-Endemic time-series model using *hhh4*, a framework developed
76 by Held, Höhle and Hofmann to study the separate impact of covariates on importation, cross-regional
77 transmission and local transmissions on aggregated case counts [14,15]. We adapted this framework
78 to daily case counts and applied it to the daily number of measles cases per department (NUTS3 levels)
79 in France reported to the European Center for Disease Prevention and Control (ECDC) between January
80 2009 and December 2018. We computed the average values of vaccine uptake and the number of cases
81 per department in the past three years to mimic the timeframe used to define the elimination status,
82 and modelled their impact on the local risks of outbreaks.

83 Methods

84 Description of the *hhh4* framework

85 We used the modelling framework implemented in the “*hhh4*” model, which is part of the R package
86 “*surveillance*” [15], to analyse infectious disease case counts. All the notations are defined in Table 1.
87 The expected number of cases ($\mu_{i,t}$) reported in the region i at time t depends on three sources of
88 transmission (called “components”):

- 89 i. The *autoregressive* component ($\lambda_{i,t}$) represents the impact of $Y_{i,t-1}$, the number of cases in i
90 at the previous time step, on the number of cases in i at t . The number of new cases expected
91 from the autoregressive component is the product of predictors $\lambda_{i,t}$ and $Y_{i,t-1}$. A high value of
92 $\lambda_{i,t}$ indicates that, if there are cases in i , there is potential for high transmission levels. On the
93 other hand, if $\lambda_{i,t}$ is low, cases in i are unlikely to lead to much local transmission.
- 94 ii. The *neighbourhood* component ($\phi_{i,t}$) represents the impact of $Y_{j,t-1}$, the number of cases
95 reported in regions around i at the previous time step, on the number of cases in i at t . The
96 exact impact of cases in these regions on cases in i is determined by a distance matrix ω which
97 quantifies the connectivity between the different regions. If $\phi_{i,t}$ is high, cases in regions around
98 i are more likely to cause new cases in i , whereas a low value of $\phi_{i,t}$ indicates that cross
99 regional transmissions towards i are less likely.
- 100 iii. The *endemic* component ($\nu_{i,t}$) represents the background number of new cases occurring in
101 region i , regardless of the current number of cases in i , or in the regions around i . If $\nu_{i,t}$ is high,
102 new cases in i are common, regardless of the number of cases in or around i at the previous
103 time step. Since the endemic component does not depend on Y_{t-1} , it represents the
104 background importations that cannot be linked to the mechanistic components. Therefore,
105 these cases either correspond to importations from outside the modelled area (France in our
106 case), or cases that are not otherwise predicted by the other two components.

107 The full equation for the expected number of cases in region i at time t is:

$$108 \quad \mu_{i,t} = v_{i,t} + \lambda_{i,t} * Y_{i,t-1} + \phi_{i,t} * \sum_{j \neq i} (\omega_{ji} * Y_{j,t-1}) \quad (1)$$

109 The predictors $\lambda_{i,t}$, $\phi_{i,t}$ and $v_{i,t}$ are independently impacted by different covariates, i.e., a covariate
 110 may be associated with a reduction of importations, but have little impact on the spread of the virus
 111 within the region. We assume that $Y_{i,t}$, the number of observed cases at t in i , follows a negative
 112 binomial distribution to allow for overdispersion [16]. The overdispersion parameter ψ is estimated.

113 The predictors $\lambda_{i,t}$, $\phi_{i,t}$ and $v_{i,t}$ are estimated using log-linear regressions. For each predictor, we
 114 estimate: i) The intercept α (identical across spatial units), and ii) the vector of coefficients β associated
 115 with $z_{i,t}$ the vector of covariates at t in i included in each component.

$$116 \quad \log(\lambda_{i,t}) = \alpha^{(\lambda)} + \beta^{(\lambda)} * z_{it}^{(\lambda)} \quad (2)$$

$$117 \quad \log(\phi_{i,t}) = \alpha^{(\phi)} + \beta^{(\phi)} * z_{it}^{(\phi)} \quad (3)$$

$$118 \quad \log(v_{i,t}) = \alpha^{(v)} + \beta^{(v)} * z_{it}^{(v)} \quad (4)$$

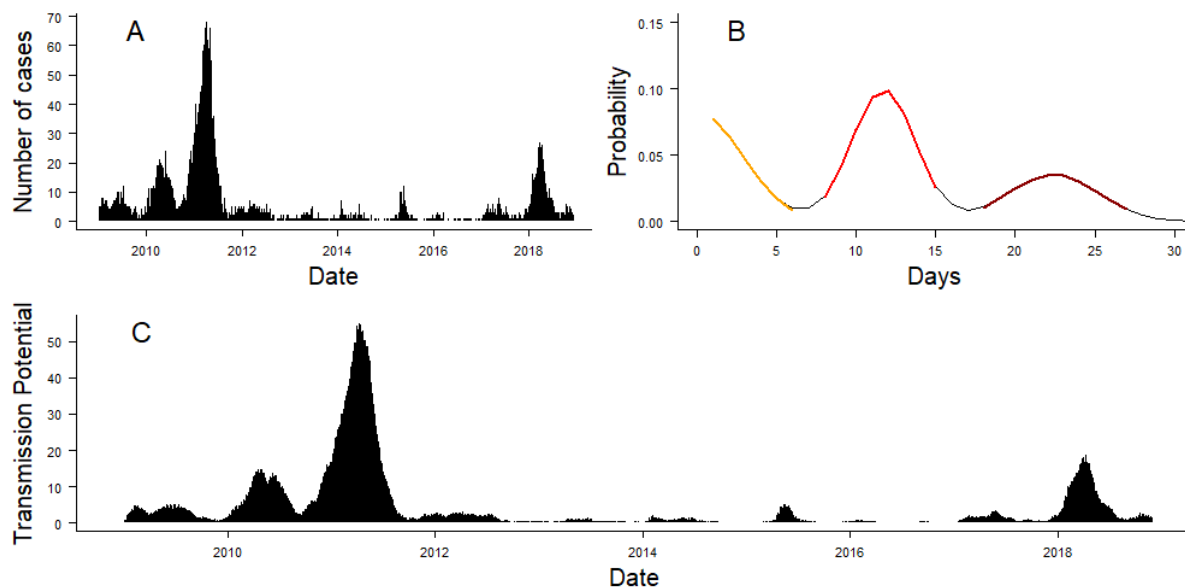
119 *Table 1: Table of notations of all variables and distributions defined in the methods.*

Parameter	Definition
i, j	Regions
t	Time
$Y_{i,t}$	Number of cases reported in the region i at time t
$Y'_{i,t}$	Potential for transmission in the region i at time t
$\mu_{i,t}$	Average number of cases predicted in the region i at time t
λ	Autoregressive predictor
ϕ	Neighbourhood predictor
v	Endemic predictor
ω	Connectivity matrix
α	Intercept
β	Vector of coefficients
z	Matrix of covariates
$f(t)$	Distribution of the serial interval
m_{it}	Number of inhabitants in the region i at time t
d_{ij}	Distance between regions i and j
γ, δ, ϵ	Parameters of the exponential gravity model

u_{it}	Average vaccine coverage in the region i at time t
n_{it}	Recent incidence per million in the region i at time t
N_{it}	Category of recent incidence in the region i at time t
s_{it}	Surface area of the region i at time t

120 Data

121 The observed case counts $Y_{i,t}$ was computed from 14,461 cases (10,988 confirmed and 3,473 probable
 122 cases) routinely collected in metropolitan France, and reported to the ECDC between January 2009 and
 123 December 2018 (Figure 1A). This data was retrieved on The European Surveillance System (TESSy) on
 124 22 January 2019. The cases were stratified by the metropolitan department they were reported in. The
 125 department correspond to French NUTS3 regions. We excluded three cases where this information was
 126 not available. We used the date of symptom onset reported for each case to compute the daily number
 127 of cases from 2009 to 2018 per department.



128

129 *Figure 1: Panel A: Daily number of cases reported in France between 1st January 2009 and 30th November 2018. Panel B:*
 130 *Distribution of the composite serial interval used in the model. The different colours of the curve correspond to the three*
 131 *scenarios used to compute the distribution of the serial interval (orange: serial interval when missing ancestor; red: serial*
 132 *interval without unreported case, brown: serial interval when the case between the two reported cases was missing). Panel C:*
 133 *Transmission potential, which was computed by convolving the number of cases in the last 30 days with the composite serial*
 134 *interval.*

135 Adaptation of hhh4 to daily case counts

136 In *hhh4*, the average number of new cases stemming from the autoregressive and neighbourhood
 137 components depends on the number of cases at the previous time step. Therefore, if we use daily case
 138 counts, the number of cases at t is only impacted by the number of cases the day before. In reality,
 139 however, the serial interval of measles is estimated to be 11 days on average [17]. Previous studies
 140 using *hhh4* relied on temporally aggregated case counts, which partially solved this problem: if the time

141 step is close to the average serial interval, cases of the same generation of transmission can be assumed
142 to be roughly grouped together in the same time point [18]. Nevertheless, studying weekly (or
143 fortnightly) aggregated cases counts does not reflect the distribution of the serial interval (i.e., it ignores
144 overlapping generations of transmission because of shorter or longer delays between primary and
145 secondary cases). This can lead to directly connected cases being grouped in the same time step, or
146 separated by more than one time step. This aggregation also ignores the potential for unreported cases,
147 which may lead to cases causing transmission two to three weeks after their onset date via an
148 intermediate, unobserved case. Finally, the starting date of aggregation influences how cases are
149 grouped, which can lead to discrepancies in the parameter estimates.

150 Recent developments in the *surveillance* package included weight estimation to represent the relative
151 impact of previous time steps on the number of cases at t [19]. Since we are using daily case counts,
152 we set the weights of the different time steps from the distribution of the serial interval. We computed
153 Y'_{it} , the transmission potential for each department and time step, by multiplying the number of recent
154 cases by the distribution of the serial interval $f(t)$: $Y'_{it} = \sum_{k=1}^{50} Y_{i,t-k} * f(k)$. Only a subset of measles
155 cases are reported to the surveillance system [20], therefore we accounted for the risks of unreported
156 cases by computing a composite serial interval from three different transmission scenarios (Figure 1B):

- 157 1- In case of direct transmission between two cases i and j , the number of days between the two
158 cases $f_1(t)$ follows a Normal distribution truncated at 0: $f_1(t) \sim N(11.7, 2)$ [17].
- 159 2- In case of unreported cases between i and j , the number of days between the two cases $f_2(t)$
160 follows a Normal distribution truncated at 0: $f_2(t) \sim N(23.4, \sqrt{8})$. This distribution corresponds
161 to the convolution of $f_1(t)$ with itself.
- 162 3- If i and j share the same unreported index case, the number of days between i and j follows a
163 half-Normal distribution (excluding 0) of standard deviation $\sqrt{8}$ days. This distribution
164 corresponds to the distribution of the difference of $f_1(t)$ with itself, excluding values below 1.
165 We added this last scenario to account for multiple concurrent importations stemming from
166 an unreported infector.

167 We considered that 50% of the composite serial interval reflected direct transmission (scenario 1,
168 without missing generations between cases), and 50% came from the two scenarios with unreported
169 cases (scenarios 2 and 3). The distribution of the composite serial interval is shown in Figure 1B. We ran
170 sensitivity analysis to estimate the parameters of the model using composite serial intervals computed
171 with different proportions of direct transmission, and observed it had little influence on the estimation
172 of each parameter (Supplement Section 1).

173 Connectivity between departments

174 In the *hhh4* framework, the average number of cases caused in the department i at time t by cases
175 from another department j is quantified by the neighbourhood component. It is equal to $\phi_{i,t} * \omega_{ji} *$
176 $Y_{j,t-1}$ (Equation 1). Therefore, the number of cases caused by cases from j in i in *hhh4* is influenced by
177 three factors:

- 178 • The susceptibility of the department i , quantified by the neighbourhood predictor $\phi_{i,t}$, defined
179 as $\log(\phi_{i,t}) = \alpha^{(\phi)} + \beta^{(\phi)} * z_{it}^{(\phi)}$.
- 180 • The number of connections from j to i , calculated using an exponential gravity model [21],
181 whereby the number of connections between i and j is proportional to the product of the
182 number of inhabitants in the department of origin m_j , the department of destination m_i and
183 an exponential decrease in the distance between i and j d_{ji} . Therefore, the number of
184 connections from j to i was calculated as $w_{ji} = e^{-\delta d_{ji}} m_{it}^\epsilon m_{jt}^\gamma$.
- 185 • The proportion of the population in j that is infectious.

186 Therefore, the average number of cases expected from department j to department i at t can be
187 written as the product of these three factors:

$$\begin{aligned} 188 \quad Y_{j,i,t} &= \exp\left(\alpha^{(\phi)} + \beta^{(\phi)} * z_{it}^{(\phi)}\right) * e^{-\delta d_{ji}} m_{it}^\epsilon m_{jt}^\gamma * \frac{Y_{j,t-1}}{m_{jt}} \\ 189 \quad &= \exp\left(\alpha^{(\phi)} + \beta^{(\phi)} * z_{it}^{(\phi)} * \epsilon * \log(m_{it})\right) * \frac{e^{-\delta d_{ji}} m_{jt}^\gamma}{m_{jt}} * Y_{j,t-1} \end{aligned}$$

190 Therefore, the log-population $\log(m_{it})$ was added as a covariate of the predictor of the
191 neighbourhood component ϕ . The number of inhabitants in each French department between 2009
192 and 2018 was taken from the INSEE website [22].

193 We implemented two models with different methods to compute the distance between departments
194 d_{ji} .

- 195 1. In Model 1, every department can be connected to each other, therefore only importations
196 coming from outside the departments included in the study fall into the endemic component.
197 The distance matrix was computed using the distance between the population centroids of
198 each department, which were calculated using the $1km^2$ European Grid dataset [23]. This
199 dataset contains the number of inhabitants in each grid cell covering the country (resolution
200 1km). We computed the weighted population centre in each department using the R function
201 *zonal* from the package raster[24] and calculated the distance between population centres.

202
$$Y_{ji,t} = \phi_{it} * e^{-\delta d_{ji}} * \frac{m_{jt}^Y}{m_{jt}} * Y_{j,t-1}$$

203 2. In Model 2, the neighbourhood component only takes into account transmission between
204 neighbouring departments, assuming that cross-regional transmissions between non-
205 neighbouring departments would be captured by the baseline number of daily importations
206 (i.e. the endemic component):

207
$$Y_{ji,t} = \begin{cases} \phi_{it} * \frac{m_{jt}^Y}{m_{jt}} * Y_{j,t-1} & \text{if } i \text{ and } j \text{ share a border} \\ 0 & \text{otherwise} \end{cases}$$

208 Therefore, the neighbourhood component in Model 1 includes both the neighbourhood component
209 and part of the endemic transmission in Model 2.

210 Covariates

211 Different covariates can be added in each component of the *hh4* framework [25]. We implemented
212 the same set of covariates in the two models. The two covariates of interest were the impact of vaccine
213 coverage and the category of incidence in each department in the past three years. We chose this
214 timeframe in order to match the requirements of the elimination status assessment. We also included
215 the number of inhabitants, the surface area of each department, and the seasonality as control
216 variables, as explained below:

217 Vaccine coverage

218 For each department i and time step t , we computed $u_{i,t}$, the average proportion unvaccinated in the
219 department i over the 3 years prior to t according to local coverage reports. We averaged over the past
220 three years in order to use the same timeframe as the elimination status assessment. We used the
221 yearly first dose uptake among 2-year-old children in each French department between 2006 and 2017.
222 This data is publicly available on the website Santé Publique France [26–28]. The uptake of the second
223 dose was not reported before 2010, and many departments had missing entries after 2010. Therefore,
224 only the local coverage of the first dose was used in the model.

225 Since 26% of the entries in the coverage dataset were missing, we ran a beta mixed model to infer the
226 missing values. We used the time and squared time (in years) as covariates, and random effects
227 stratified by department. We used the average prediction to infer the missing values from the fitted
228 model and get the complete vaccine coverage dataset. More details on the regression, and the
229 sensitivity analyses that were run are presented in the Appendix (Supplement Section 2). All values of

230 coverage in 2009 were missing, and were not imputed; we computed the average vaccine coverage in
231 2010, 2011, and 2012 using only two of the three previous years.

232 Adding the log-proportion of unvaccinated to the model was the most appropriate approach, since it
233 allows the rate of disease spread (i.e. the value of the predictors λ , ν , and ϕ) to be proportional to the
234 density of susceptibles [25]. Therefore, we calculated the average log-proportion of unvaccinated in
235 the three years before t and added it as a covariate in all three components.

236 Impact of recent incidence

237 This covariate quantifies the impact of past outbreaks on current transmission. Departments are eligible
238 for WHO certification of elimination status if they have maintained low levels of transmission over the
239 past three years [7]. Therefore, we computed $n_{i,t}$, the number of cases per million reported between
240 a month and three years before t in i . We excluded cases reported in the last month since recent cases
241 may be directly linked to current transmission.

$$242 \quad n_{i,t} = 1,000,000 * \sum_{\substack{T < t - 365 * 3 \\ T > (t - 30)}} \frac{Y_{it}}{m_{it}}$$

243 We aggregated $n_{i,t}$ in three categories: i) $N_{i,t}^{(0)} = \begin{cases} 1 & \text{if } n_{i,t} < 10 \\ 0 & \text{otherwise} \end{cases}$: very limited transmission in recent
244 years, department potentially eligible for elimination (30% of entries) ; ii) $N_{i,t}^{(1)} =$
245 $\begin{cases} 1 & \text{if } 10 \leq n_{i,t} < 45 \\ 0 & \text{otherwise} \end{cases}$: Moderate transmission in recent years (36% of entries); iii) $N_{i,t}^{(2)} =$
246 $\begin{cases} 1 & \text{if } n_{i,t} \geq 45 \\ 0 & \text{otherwise} \end{cases}$: major outbreak reported in the department in recent years. The threshold of 45 cases
247 per million corresponds to the last tercile of $n_{i,t}$, hence 33% of $n_{i,t}$ fall into this last category.

248 Computing the level of recent incidence required the number of cases per department in the past three
249 years. Therefore, since this analysis integrates case counts data from 2009, we needed to compute the
250 incidence in each department between 2006 and 2008. Less than 50 cases were reported in France per
251 year in 2006 and 2007 [29], therefore we considered their contribution to the recent level of incidence
252 per department was null. On the other hand, 597 measles cases were reported to the ECDC in France
253 in 2008, but were not stratified by department. Therefore, we used the number of cases reported per
254 department in 2008 on Sante-Publique-France (597 cases overall, mostly reported in the second half of
255 2008 [30]) and integrated them in the computation of $N_{i,t}$ for $t < 2012$.

256 The level of recent incidence was a covariate in all three components.

257 Number of inhabitants and surface area

258 In the subsection “Connectivity between departments”, we discussed the impact of the number of
259 inhabitants on the number of movements between departments. Furthermore, several studies have
260 indicated a potential association between the population density and the number of secondary
261 transmissions [31–33]. Therefore, we controlled for the impact of the number of inhabitants in each
262 department, and the surface area (i.e., the geographical size) on the number of local transmissions.

263 The log-number of inhabitants $\log(m_{i,t})$ in the department i at time t was added as a covariate in all
264 three components. The log-surface of the department $\log(s_{i,t})$ was added as a covariate in the
265 autoregressive component.

266 Seasonality

267 We control for the impact of the seasonality of measles outbreaks in France on transmission by adding
268 two covariates (sine-cosine) to all three components.

269 Full model equations for predictors

270 The covariates are all integrated in the covariate vectors in the equations 2, 3 and 4, yielding:

271 Autoregressive predictor: $\beta^{(\lambda)} z_{it}^{(\lambda)} = \beta_u^{(\lambda)} \log(u_{i,t}) + \beta_{N^{(1)}}^{(\lambda)} N_{i,t}^{(1)} + \beta_{N^{(2)}}^{(\lambda)} N_{i,t}^{(2)} + \beta_m^{(\lambda)} \log(m_{i,t}) +$
272 $\beta_s^{(\lambda)} \log(s_{i,t}) + \beta_{\cos}^{(\lambda)} \cos\left(\frac{2\pi t}{365}\right) + \beta_{\sin}^{(\lambda)} \sin\left(\frac{2\pi t}{365}\right)$

273 Neighbourhood predictor: $\beta^{(\phi)} z_{it}^{(\phi)} = \beta_u^{(\phi)} \log(u_{i,t}) + \beta_{N^{(1)}}^{(\phi)} N_{i,t}^{(1)} + \beta_{N^{(2)}}^{(\phi)} N_{i,t}^{(2)} + \beta_m^{(\phi)} \log(m_{i,t}) +$
274 $\beta_{\cos}^{(\phi)} \cos\left(\frac{2\pi t}{365}\right) + \beta_{\sin}^{(\phi)} \sin\left(\frac{2\pi t}{365}\right)$

275 Endemic predictor: $\beta^{(v)} z_{it}^{(v)} = \beta_u^{(v)} \log(u_{i,t}) + \beta_{N^{(1)}}^{(v)} N_{i,t}^{(1)} + \beta_{N^{(2)}}^{(v)} N_{i,t}^{(2)} + \beta_m^{(v)} \log(m_{i,t}) +$
276 $\beta_{\cos}^{(v)} \cos\left(\frac{2\pi t}{365}\right) + \beta_{\sin}^{(v)} \sin\left(\frac{2\pi t}{365}\right)$.

277 Model calibration

278 A model is deemed well-calibrated if it is able to correctly identify its own uncertainty in making
279 predictions [34]. The most straightforward method to evaluate whether *hhh4* models are well-
280 calibrated is to generate a one-step-ahead forecast over a chosen test period and compare them with
281 the data [15]. Since we use daily case counts, this method would only assess the ability of the models
282 to capture the number of cases on the next day. We explored the calibration of our models several days
283 ahead. To do so, we selected the last two years of data as the test period, fit the model up to each day,
284 and simulated the number of cases over the next 3, 7, 10 and 14 days for each day of the test period in
285 each department. For each date, we ran at least 100,000 simulations. If the number of cases observed
286 in the data had not been generated in 100,000 simulations, we ran simulations until it was reached.

287 From these simulations, we generated the predictive probability distribution at each time step in each
288 department. In a model with perfect calibration, the actual number of cases follows the predictive
289 probability distribution ($\mu_{it} \sim P_{it}$ for all predictive distributions P_{it}), i.e., the probability integral
290 transform (PIT) histogram is uniform. We computed the PIT histograms in both models for predictions
291 over 3, 7, 10, and 14 days. The PIT histograms were computed using a non-randomised yet uniform
292 version of the PIT histogram correcting for the use of discrete values described in Czado et al [35] and
293 implemented in *hhh4*.

294 The PIT histograms were used to estimate whether the short-term forecasts were in line with the data,
295 and whether the models were consistently missing some scenarios of transmission.

296 [Simulation study](#)

297 In order to highlight the impact of variations in the local vaccine coverage or the level of recent
298 transmission on the risks of outbreaks, we generated simulations of the number of cases in France
299 across one year under different conditions. To compute these simulations, we used the last values of
300 average vaccine coverage (the average was computed from the values in 2015, 2016, and 2017) and
301 the levels of recent incidence in mid-2018, and simulated the daily number of cases between the 1st of
302 August 2018 and the 31st of December 2019. We started the simulations during the period of the year
303 associated with the lowest number of cases (i.e., on the 1st of August), in order to avoid biases. Indeed,
304 if we had used the last three months of data (until November 2018), some departments may have been
305 repeatedly associated with higher numbers of cases in our simulations, not because they are more at
306 risk of importation or transmission, but because there had been cases reported in these departments
307 at the beginning of the epidemic year. We were only interested in highlighting the impact of variations
308 in coverage and recent transmission, rather than predicting the level of transmission for the entire year
309 of 2019.

310 We generated 100 samples of the regression coefficients using the variance-covariance matrix and
311 assumed they followed a multivariate normal distribution. For each sample, we computed the values
312 of the three predictors between the 1st of August 2018 and the 31st of December 2019, and simulated
313 the daily number of cases in each department across the year. We ran 100 simulations per sample (i.e.
314 10,000 simulations were generated per scenario).

315 We studied four scenarios: i) Using the latest local values of coverage (averaged over the past three
316 years), population and category of recent incidence, ii) Increasing the vaccination coverage in each
317 department by three percent, iii) Decreasing the vaccination coverage in each department by three
318 percent, and iv) setting the recent incidence in each department to minimal levels (i.e. conditions
319 fulfilling the WHO elimination status requirements).

320 Finally, since tourism and local events can lead to mass gatherings and trigger repeated importations
321 independent of parameters included in the model [36,37], we studied the impact of repeated local
322 importations of cases into specific departments. To do so, we simulated one year of transmission (i.e.,
323 until the end of 2019) following the importations of 10 cases in a given department in December 2018.
324 In these simulations, we did not allow for any other baseline importations throughout the year, in order
325 to assess the potential for geographical spread throughout the country after importation in one
326 department.

327 Results

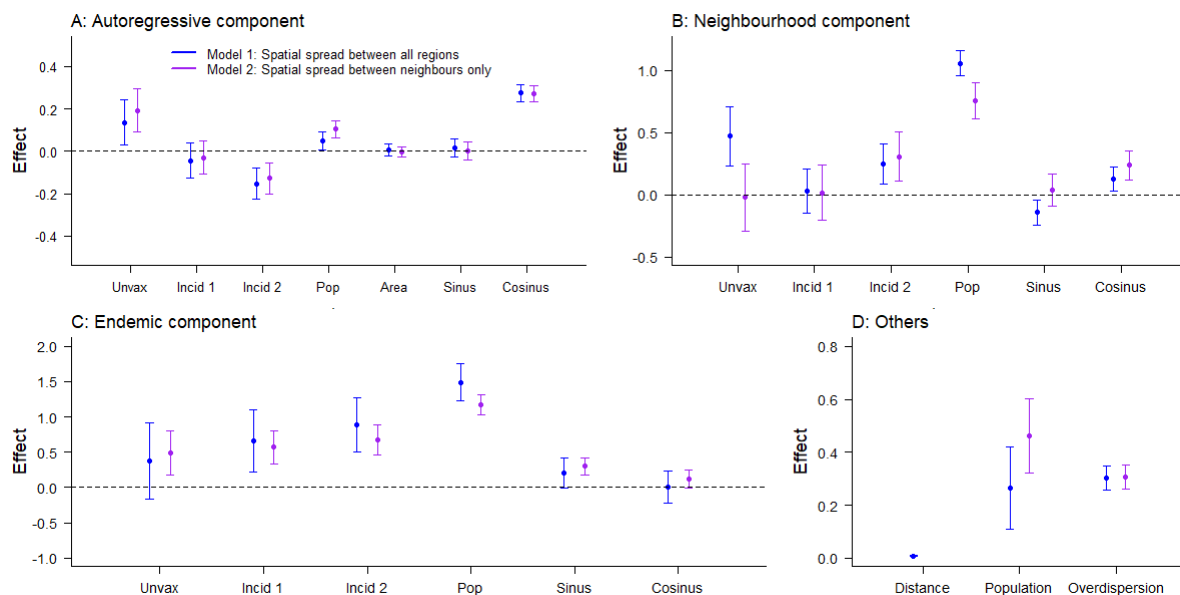
328 Impact of the covariates on each component

329 The parameter estimates obtained in both models are shown in Figure 2. Values above 0 show
330 aggravating effects associated with an increase in the number of expected cases at the next time step.
331 For both models, departments with a high proportion unvaccinated in the past three years were
332 associated with a higher number of expected cases in the autoregressive (Model 1: 0.14 [0.03 - 0.24] ;
333 Model 2: 0.19 [0.09 - 0.29]) and the endemic component (Model 1: 0.37 [-0.17 - 0.91] ; Model 2: 0.48
334 [0.17 - 0.80]). This indicates that these departments were at higher risks of background importations,
335 and secondary transmission upon importation. In both components, the effect of vaccination was
336 slightly stronger in Model 2, where cross-regional transmission is restricted to neighbouring
337 departments, than in Model 1, where cross-regional transmission can happen between all
338 departments, although the confidence intervals overlapped. In Model 1, the proportion unvaccinated
339 also had an aggravating effect on the number of cross-departmental transmissions (0.47 [0.23 - 0.71]),
340 whereas in Model 2 there was no clear association between the proportion unvaccinated and an
341 increase in cross-regional transmission (-0.02 [-0.29 - 0.25]). The differences between the models'
342 coefficients were due to the cross-regional transmission in Model 1 corresponding to both the
343 neighbourhood component and some of the endemic transmission in Model 2.

344 The association between the level of incidence over the past three years (parameters: *immun 1* and
345 *immun 2* in Figure 2) and the components of transmission was similar in both models. In the auto-
346 regressive component, departments that reported high incidence over the past three years (*immun 2*)
347 were associated with fewer secondary cases per case in the department (Model 1: -0.15 [-0.23 - -0.08];
348 Model 2: -0.13 [-0.20 - -0.06]). This could be linked to outbreak-induced immunity causing a depletion
349 of susceptibles in departments where incidence was high over the past few years. On the other hand,
350 the parameters associated with *immun 2* were above 0 in the neighbourhood and endemic
351 components, which indicates that departments with high incidence in the past three years were more
352 at risk of cross-regional transmission and background importations (Model 1: Endemic 0.89 [0.50 –

353 1.27]; Neighbourhood: 0.25 [0.09 – 0.41]; Model 2: Endemic 0.67 [0.46 – 0.89]; Neighbourhood: 0.31
 354 [0.11 – 0.51]). The parameter *immun 1* was only significantly different from 0 in the endemic
 355 component (Model 1: 0.66 [0.22 – 1.10]; Model 2: 0.57 [0.34 – 0.80]), meaning departments that
 356 recently reported moderate levels of transmission were associated with more background
 357 importations, but no difference was noticeable in cross-regional or within-region transmission.

358 The other covariates included in the model showed that the number of inhabitants in a department
 359 had an important impact on both the endemic and neighbourhood components: departments with
 360 more individuals were more likely to report background importations and cross-regional transmission.
 361 On the other hand, the population and the surface area of the departments had no impact on the
 362 autoregressive component. We also observed a strong impact of seasonality on the three components
 363 (Figure 2). Indeed, the peak values of the predictors were 20 to 37% higher than the average value in
 364 all components of transmission (Supplement Section 3). The peak of the autoregressive component
 365 was in February for both models, the endemic peak was in May for Model 1 (April in Model 2), whereas
 366 the neighbourhood component peaked in December in Model 1 (March in Model 2).



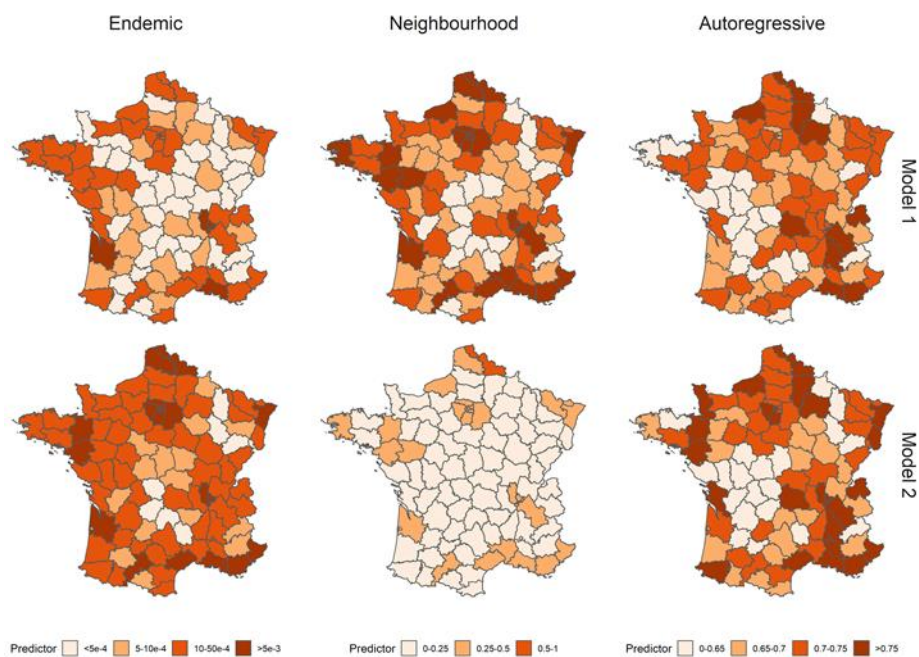
367

368 *Figure 2: Estimates of the parameters in each component of Model 1 (blue) and Model 2 (purple): Panel A: Autoregressive*
 369 *component; Panel B: Neighbourhood component; Panel C: Endemic component; Panel D: Other coefficients. The y-axis. unvax*
 370 *corresponds to the effect of $u_{i,t}$, the mean proportion unvaccinated over the three years before t in i ; incid1 and incid2*
 371 *correspond to the effect of $N_{i,t}^1$ and $N_{i,t}^2$, the category of incidence in the three years before t in i ; pop corresponds to the effect*
 372 *of $m_{i,t}$, the number of inhabitants at t in i ; area corresponds to the effect of the surface; sin and cos correspond to the effects*
 373 *of seasonality; distance and population correspond to the spatial parameters of the connectivity matrix w (δ and γ);*
 374 *overdisp is the estimate of the log-overdispersion parameter in the negative binomial distribution of $Y_{i,t}$. Dots show the mean*
 375 *values associated with the parameters; arrows show the 95% Confidence interval. Note different y-axes between graphs.*

376 Using the mean parameter estimates, and the latest values of vaccination coverage, incidence, and
 377 number of inhabitants per department, we computed the local predictors ϕ_i , λ_i , and v_i in both models
 378 to highlight the spatial heterogeneity of the transmission risks (Figure 3). The predictors were computed

379 ignoring the impact of seasonality, which does not change the geographic distribution of risks since it
380 is not region-dependent in the models. Therefore, the maps correspond to the average local value of
381 the predictors the year following the last data entry (i.e. the 30th of November 2018). The geographic
382 distributions of the autoregressive predictor are similar in Model 1 and Model 2. This indicates that the
383 same departments were classified as having higher risks of local transmission in both models. Areas
384 with lower values of vaccine uptake such as the South East and South West of France were associated
385 with higher risks of secondary transmission. Indeed, the highest values of within-region transmission
386 were reported in Bouches-du-Rhône and Var (in the South East of France). Populous departments in
387 the North of France were also at risk of secondary transmission despite higher vaccination coverage.

388 As expected, the overall number of baseline importations in Model 1 was lower than in Model 2, which
389 was compensated by a higher number of cross-regional transmissions (Figure 3). This shows that some
390 of the cases that could not be linked to local transmission, or transmission between neighbouring
391 departments in Model 2, were classified as cross-regional transmissions in Model 1, which would
392 indicate long-distance transmission events. In both models, departments with a higher number of
393 inhabitants were most at-risk of cross-regional and baseline importations, which corresponds to the
394 strong association between the number of inhabitants and the endemic and neighbourhood
395 components highlighted in Figure 2. Departments like Bouches-du-Rhône that combine a high number
396 of inhabitants with low vaccine coverage were associated with the highest number of baseline and
397 cross-regional importations in both models. The variations in the autoregressive component were
398 smaller than in the importation-related components: For instance, the highest autoregressive predictor
399 value (Var: 0.81 [0.74 - 0.88]) was 35% higher than the lowest value (Lozère: 0.60 [0.53 – 0.66]) in Model
400 1, whereas the number of baseline importations in Bouches-du-Rhône was more than 100 times above
401 the number of importations in Lozère (South of France). This can be explained by the coefficients of the
402 autoregressive components being much closer to 0 than the most extreme coefficients in the
403 importation-related components (Figure 2).



404

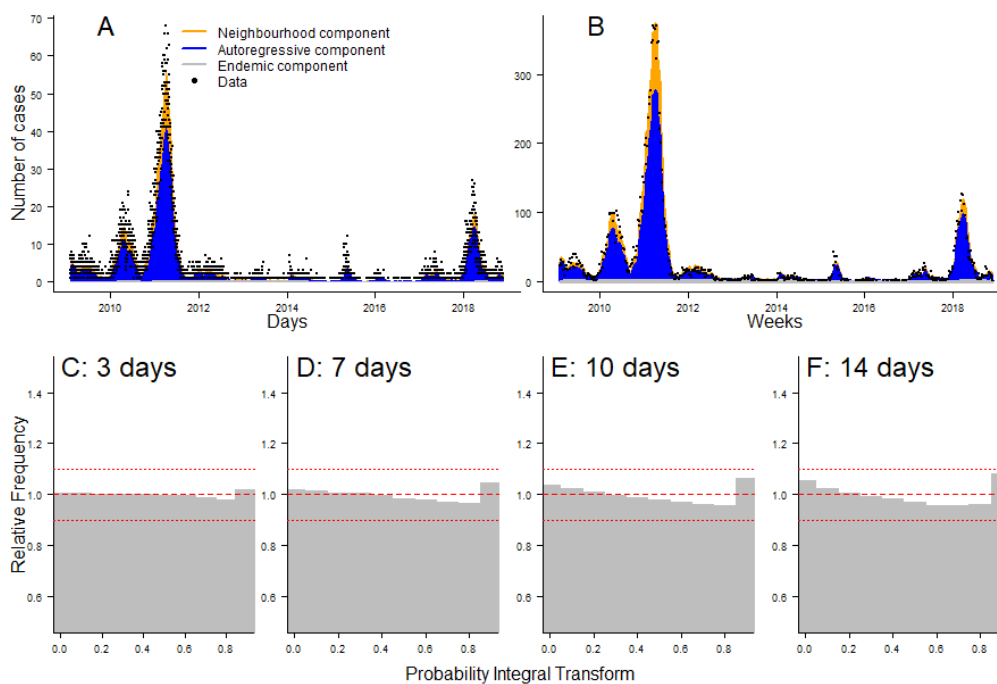
405 *Figure 3: Average values of the endemic, neighbourhood, and autoregressive predictors per department in Model 1 (upper row)*
406 *and Model 2 (lower row) over the year 2019. Since the absolute values are expected to vary over the year because of seasonality,*
407 *the panels show the relative geographical heterogeneity. The endemic predictor corresponds to the number of importations*
408 *per day per department, whereas the autoregressive predictor corresponds to the number of secondary cases per case in each*
409 *department. The absolute value of the neighbourhood predictor is harder to interpret directly since it is multiplied by the*
410 *connectivity matrix in the equation. Higher values were associated with departments with higher risks of observing cases*
411 *following population movements.*

412 Model fit and calibration

413 The daily and weekly fits of Model 1 and Model 2 indicate that they were able to match the transmission
414 dynamics observed in France between 2009 and 2017, despite wide variations in the annual number of
415 cases (Figure 4 Panel A and B, Supplement Section 4). In years where active transmission was reported,
416 most of the cases stemmed from the autoregressive component, indicating that the local outbreaks
417 were sustained by transmission within the departments. Indeed, across all years, the autoregressive
418 component accounted for 72.9% of the cases, whereas 23.7% of the cases came from cross-regional
419 transmission, and 3.4% from the endemic component (Supplement Figure S12). This shows that in
420 Model 1, 97.6% of the cases were explained by the transmission stemming from other cases reported
421 in the dataset (93.2% in Model 2). The endemic component described the minority of isolated cases
422 that could not be linked to any concurrent transmission cluster. Therefore, these cases would be more
423 likely to be reported at times of low national levels of transmission when no other case could be linked
424 to them, which explains the shift in seasonality of the endemic component observed in Figure 2 and
425 Supplement Section 3.

426 In order to visually assess the calibration of the model, and its ability to provide reliable short-term
427 predictions for the number of cases per department, we generated PIT histograms showing the

428 probability integral transform obtained when forecasting the number of cases 3, 7, 10, and 14 days
429 ahead (Figure 4, Panels C to F). The PIT histogram is uniform for predictions 3 and 7 days ahead (all
430 groups are above 0.9 and below 1.1), which shows the number of occurrences where the predictions
431 of the model did not capture the number of cases one week ahead was not higher than expected under
432 a uniform distribution. As we increased the number of days of forecast, there were more occurrences
433 of the model mis-predicting the number of cases to come. Indeed, the U-shape observed in Panel F of
434 Figure 4 indicates the model was less capable of identifying extreme events two weeks in advance. The
435 calibration study indicated that Model 2 was more prone to under-estimating the number of cases than
436 Model 1, and showed signs of bias for the 7, 10, and 14-day predictions (Supplement Section 4). The
437 national number of cases predicted by Model 1 and Model 2 were similar, and match the data for
438 predictions 7 days ahead (Supplement Figure S11). The AIC scores and the calibration study indicated
439 Model 1 was able to fit the data better than Model 2 and was better calibrated. The rest of the Results
440 section therefore focuses on the conclusions reached using Model 1. The equivalent analysis run on
441 Model 2 is presented in the Supplementary Section 4.



442

443 *Figure 4: Panel A and B: Daily and weekly fit between the data and Model 1. The inferred number of cases is split among the*
444 *three components of the model. Panel C to F: PIT histograms of Model 1, generated respectively for predictions 3, 7, 10, and*
445 *14 days ahead.*

446 [Impact of vaccination and recent incidence on onwards transmission](#)

447 In order to illustrate the impact of recent outbreaks and variations in vaccine coverage on the
448 transmission dynamics in France, we generated 10,000 simulations and computed the number of cases
449 per department in 2019. We ran the simulation from August 2018 (during the historically low

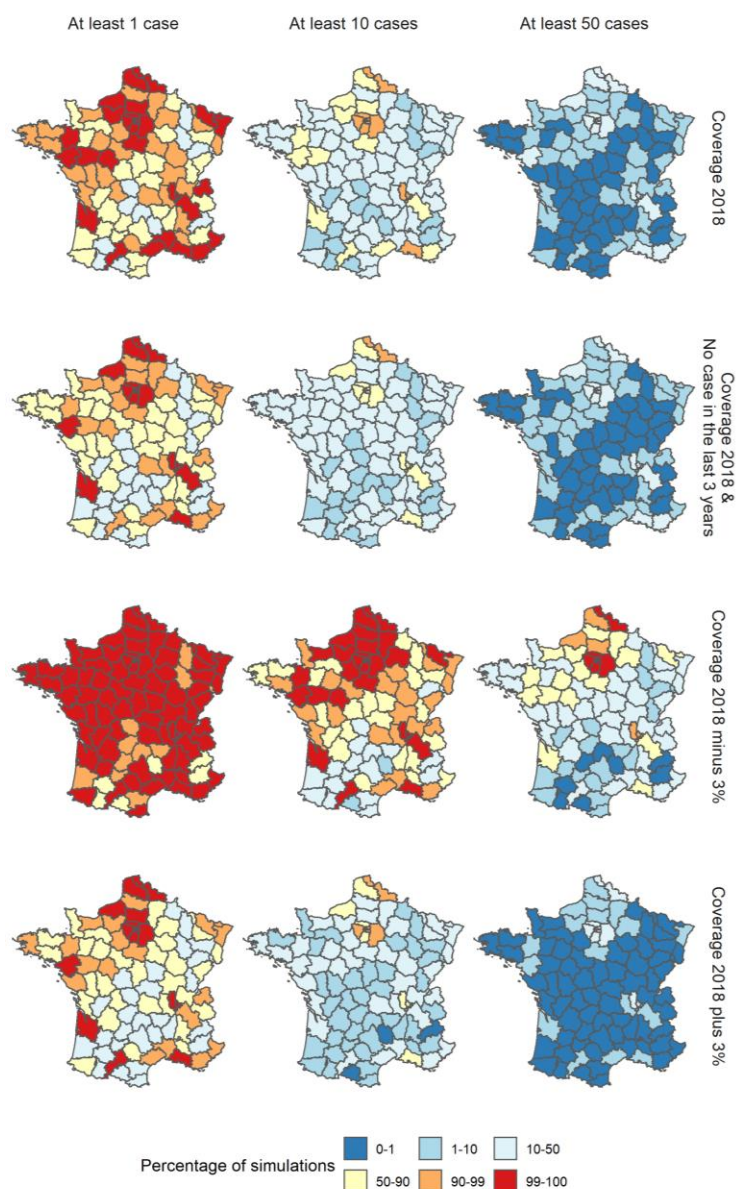
450 transmission season), until 31st December 2019. We generated four sets of simulations under different
451 initial conditions: using the last measures of average local vaccine coverage, category of recent
452 incidence, and number of inhabitants; increasing or decreasing the vaccine coverage by three percent,
453 and setting the category of recent incidence to 0 in each department.

454 Under the latest measures of coverage and incidence, the simulated outbreaks display a wide variation
455 in the number of cases in 2019 (minimum 100 cases, median 1,100 cases, maximum 11,100 cases).
456 Active transmission was generated in a wide range of departments. Indeed, across the simulation set,
457 44 of the 94 French departments reported more than 10 cases in at least 25% of the simulations. There
458 was noteworthy spatial heterogeneity in the levels of incidence. Indeed, in 12 departments, there was
459 no case generated in more than half of the simulations (Figure 5, top right panel). The departments
460 most vulnerable to active transmissions were highly populated urban areas, such as Paris, the Bouches-
461 du-Rhône, and the North of France. Because they are highly populated, these departments were
462 susceptible to repeated importations (they reported at least 1 case in more than 95% of the
463 simulations), which could then cause large transmission clusters. This was especially evident in the
464 South-East of France, where we highlighted that the number of secondary cases per case in the
465 department was among the highest in the country (Figure 3 and Figure 5). Numerous departments
466 were affected by large outbreaks in a subset of the simulated datasets: 27 departments reported more
467 than 50 cases in at least 5% of the simulations (Figure 5). Further, at least one major outbreak was
468 generated in the majority of the simulations: in 55% of the simulations, one department reported more
469 than 100 cases (the most commonly affected department were Paris and its surroundings, the Nord,
470 and Bouches-du-Rhône).

471 Decreasing the average three-year vaccine coverage by three percent led to an important increase in
472 the number of cases per outbreak (median 4,900 cases, more than 95% of the simulations resulted in
473 more than 1,000 cases). This was first due to an increase in importations and cross-regional
474 transmission: all 94 departments had at least one case in more than half of the simulations, 77 in at
475 least 90% of the simulations. Furthermore, the decrease in vaccination coverage resulted in higher
476 chances of uncontrolled transmissions in many departments (Figure 5, third row). On the other hand,
477 increasing the vaccine coverage by three percent caused an important drop in the number of cases
478 (median 605 cases, 80% of the simulations generated less than 1,000 cases), caused by both a decrease
479 in the number of importations, and in the potential for secondary transmission following importations.
480 Although outbreaks were still punctually generated, these events are much rarer than in the other two
481 simulation sets: in 25.8% of the simulations, at least one department generated more than 100 cases
482 (54.1% with the baseline scenario, 95.4% when we reduced the local vaccine coverage).

483 Finally, setting the local recent incidence to the minimum level in each department, which would fulfil
484 the elimination guidelines, had two opposite effects: it led to a decrease in the number of importations
485 and cross-regional transmission, and an increase in the number of infections within each department
486 (Figure 2). In this simulation set, the number of departments where no cases were generated in more
487 than half of the simulations was similar to when the vaccine coverage was increased (24 departments
488 in this simulation set, 29 when the vaccine coverage was increased, Figure 5), which shows the
489 reduction in the number of cross-regional transmission and background importations. Conversely, the
490 number of large outbreaks was only marginally inferior to the reference simulation set: in 44% of the
491 simulations, there were more than 100 cases generated in at least one department (54% in the
492 reference dataset). The geographical distribution of the risks of large outbreaks was almost identical to
493 the reference simulation set (Figure 5). Therefore, although the number of importations was reduced,
494 changing the level of recent incidence did not have a clear impact on the risks of active transmission.
495 More departments became vulnerable to secondary transmission, and despite importations in these
496 departments being rarer, they were more likely to lead to large outbreaks when they happened. The
497 two opposing effects recent incidence had on importation and transmission therefore created a
498 different dynamic of transmission observed in the simulation set, without strongly reducing the risks of
499 outbreaks.

500 Each of these simulation sets highlighted the wide range of scenarios that could be generated using the
501 parameter distributions inferred by our model. In order to gain more understanding on the spatial
502 spread and consequences of importations, we then explored the impact of localised repeated
503 importations on overall transmission.



504

505 *Figure 5: Percentage of simulations where the number of cases reported in each department in 2019 was at least 1, 10, and 50*
 506 *cases for each scenario using parameter estimates from Model 1. Each row corresponds to a different scenario: i) Reference,*
 507 *ii) Minimum level of recent incidence in each department, iii) Local vaccine coverage decreased by three percent in each*
 508 *department, iv) Local vaccine coverage increased by three percent in each department.*

509 Impact of local clusters of transmission

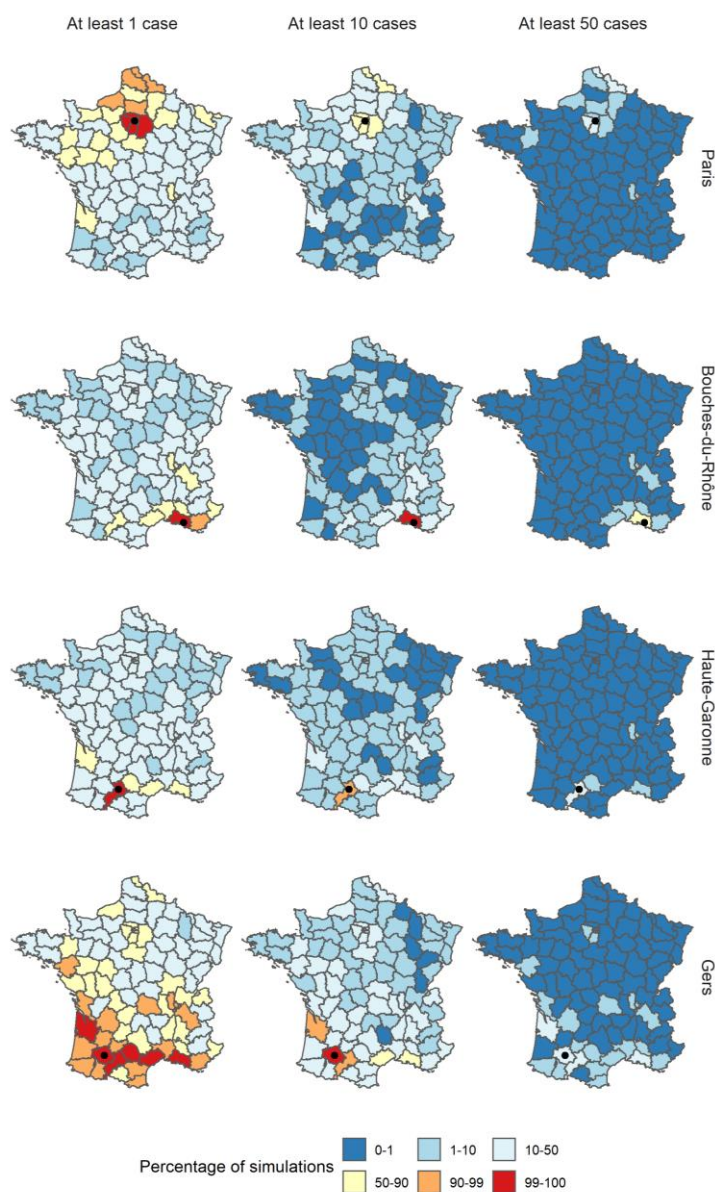
510 Since the endemic component, which can be interpreted as external importations, represented a
 511 minority of the cases in our model (Supplement Figure S12), repeated importations in a given
 512 department over a short timespan rarely occurred in the simulations. Furthermore, due to the
 513 seasonality of the endemic component, fewer importations are generated early in December to
 514 February, which corresponds to the peak period of the other components, and would therefore be
 515 more likely to cause secondary transmissions (Supplement Section 3). We simulated one year of
 516 transmission following ten importations in December 2018 to illustrate: i) the potential for local

517 outbreaks, and ii) the spatial spread of transmission following repeated local importations. We selected
518 four departments to compare the impact of repeated importations in a range of settings: Paris (many
519 inhabitants, 91% vaccine coverage, surrounded by urban areas), Bouches-du-Rhône (many inhabitants,
520 84% vaccine coverage), Haute Garonne (many inhabitants, 91% vaccine coverage but high levels of
521 recent incidence, surrounded by rural areas with lower vaccine coverage), and Gers (Rural area, 79%
522 vaccine coverage) (Figure 6).

523 Firstly, major local outbreaks in the department of importation were generated in all four simulation
524 sets, and especially in Paris and Bouches-du-Rhône, where the proportion of simulations that yielded
525 more than 100 subsequent cases in the department was 40% and 39%, respectively. In the Bouches-
526 du-Rhône, large outbreaks were mostly due to the low vaccination coverage, whereas in Paris,
527 outbreaks were mostly linked to the connectivity to nearby areas and the high number of inhabitants,
528 which meant the department was likely to attract cross-regional transmissions. Major local outbreaks
529 were rarer in the other two scenarios (9% of simulations above 100 in Haute Garonne, 10% in Gers).
530 The lower proportion of large outbreaks resulted from different factors: recent large outbreaks in Haute
531 Garonne reduced the autoregressive predictor, lowering the number of secondary cases per case
532 imported; whereas since Gers is a rural department, with a low number of inhabitants, almost all the
533 local cases were due to local transmission (auto-regressive component), with very few cross-regional
534 transmissions into Gers.

535 Conversely, the simulations where cases were imported in Gers yielded the largest spatial spread
536 throughout the country: the median number of departments that reported at least 1 case was 53 (16
537 when the importations were generated in Haute Garonne; 15 in Bouches-du-Rhône; 39 in Paris). As
538 stated in the method, the number of cross-regional transmissions is the product of the predictor and
539 the connectivity matrix, divided by the number of inhabitants in the department of origin, to represent
540 that only a fraction of commuters will be infected. Therefore, populous areas are more likely to attract
541 cross-regional transmissions, whereas more rural departments are more likely to seed outbreaks in
542 other areas. The relatively high spatial spread when cases were imported in Paris is due to the short
543 distance between Paris and its suburbs, which is then more likely to cause cross-regional transmission
544 in the northern departments. Despite the cross-regional spread observed in both of these simulations
545 sets, outbreaks remained local, and occurrences of nation-wide outbreaks were almost null. The
546 departments most at risk of outbreak following cross-regional spread were some of the direct
547 neighbours of the department of importations, or the large urban areas (Figure 6). To further explore
548 this, we ran the same simulations decreasing the vaccine coverage by three percent, which greatly
549 increased the number of departments exposed in each simulation set, and increased the risk of local
550 transmission (Supplement Section 6). Therefore, although repeated importations could cause active

551 transmission in and around the departments of importation, the current values of vaccine coverage and
552 the seasonality of transmission were able to prevent nationwide transmission.



553

554 *Figure 6: Percentage of simulations where the number of cases reported in each department in 2019 was at least 1, 10, and 50*
555 *cases following the importations of ten cases in December 2018, and using the parameter estimates from Model 1. For each*
556 *row, the department of importation is indicated by a black dot.*

557 Discussion

558 This analysis explored which local factors were associated with high risks of transmission in France over
559 the last decade. Since 2017, immunity gaps, caused by failures to vaccinate, have been linked to a
560 resurgence of measles in all WHO regions [38]. In countries near-elimination, large outbreaks have been
561 linked to heterogeneity in the levels of immunity, with pockets of susceptibles fuelling punctual
562 outbreaks despite high national vaccine uptake [1,2,4,25]. Our study showed that local values of vaccine

563 coverage were linked to lower transmission, whereas lower levels of recent incidence were not
564 associated with lower risks of local transmission. Furthermore, we highlighted that a drop of 3% in the
565 three-year vaccine coverage triggered a five-fold increase in the number of cases simulated in a year.

566 The fact that higher vaccine coverage was associated with a lower number of secondary cases is
567 consistent with prior expectations, and would confirm that the local values of first dose vaccine
568 coverage are a good indicator of the actual immunity in the population and risks of future transmission.
569 Reporting accurate values of local vaccine coverage is challenging, for instance because the vaccination
570 status of people moving regions can be hard to track and lead to measurement errors. Furthermore,
571 we did not have access to complete data on the coverage of the second MMR dose, which would be a
572 better indicator of vulnerable areas. Therefore, detecting the association between recent vaccine
573 uptake and incidence is encouraging. The impact of local vaccination coverage on transmission may
574 also be muddled by sub-regional vaccine heterogeneity. For instance, pockets of susceptibles within a
575 region, i.e. areas within the region where the vaccine coverage is substantially lower than the regional
576 average, may be at high risk of transmission and would not be observable in regional coverage [39].
577 This phenomenon can only be hypothesised here, and could be explored using local data on incidence
578 and vaccine uptake at a sub-regional scale.

579 Variations in vaccine coverage had a noticeable impact on the number of cases generated in the
580 simulation study. We showed the effects of a three percent increase and decrease of the three-year
581 average vaccine coverage on the number of cases, which highlighted the risks of uncontrolled
582 transmission in the event of a decrease of vaccine-induced protection. Events such as the disruption
583 caused by the SARS-COV-19 pandemic on routine measles vaccination campaigns could therefore highly
584 increase the risks of uncontrolled measles transmission in the years to come [40,41].

585 The departments that reported few cases per million in the past three years were associated with
586 higher risks of local transmission (autoregressive component). Therefore, according to our model,
587 regions eligible for elimination status were not associated with lower risks of onwards transmission.
588 Conversely, high levels of recent transmission were associated with a lower number of cross-regional
589 transmissions and importations, although we cannot methodologically establish the causality of this
590 association. The impact on the simulation study was clear: when we set the category of recent incidence
591 to the lowest level, departments were less exposed to cases, and spatial spread was rarer, whilst there
592 was little change in the risks of major outbreaks. The simulations showed an 'all-or-nothing' situation:
593 departments tended to report very few to no cases, whilst also being more likely to be affected by
594 outbreaks. These results would indicate that looking into the level of incidence to quantify the future

595 risks of outbreaks can be deceptive, and importations in a department with low recent incidence would
596 result in large transmission clusters.

597 We proposed a new framing of the Epidemic-Endemic model implemented in *hhh4* by adapting it to
598 daily count data using the distribution of the serial interval to compute the local transmission potential.
599 Using daily case counts allowed us to avoid biases associated with aggregated case counts, such as the
600 influence of the arbitrary aggregation date, by accounting for the impact of variation in the serial
601 intervals. We also accounted for the risks of unreported cases by computing a composite multimodal
602 serial interval, thus allowing for transmission with a missing generation, or an unreported ancestor. The
603 model was able to capture the dynamic of transmission better than the 10-day aggregated model, as
604 shown by the calibration study (Supplement Section 7). Nevertheless, our framing of the *hhh4* model
605 introduced new biases: we used a distribution of the serial interval based on previous studies rather
606 than estimating the weights during the fitting procedure and set the proportion of missing generations
607 in the composite serial interval. We explored the impact of the proportion of missing generations by
608 fitting the model with different composite serial intervals and concluded that the impact of each
609 covariate was robust to these changes (Supplement Section 1). We also integrated a potential day-of-
610 the-week effect, and observed that although it had an impact on the auto-regressive component, it did
611 not change the estimates of the other parameters, and therefore did not change the conclusions of the
612 study (Supplement Section 8).

613 Using the *hhh4* model allowed us to analyse the different impact of various covariates on local and
614 cross-regional transmission, and background importation of cases. According to the models we
615 implemented, an overwhelming majority (>90%) of the transmission came from the cross-regional and
616 local components of the regression. This indicates that in the models, the endemic component only
617 corresponds to rare background cases that could not be linked to concurrent transmission events. This
618 could point towards model misspecifications, for example, connecting unrelated importations to
619 concurrent local transmission. Since endemic transmission tends to refer to cases otherwise
620 unexplained by the mechanistic components, the seasonality of the endemic component is decoupled
621 from the other components, i.e. endemic cases are likely when local and cross-regional transmission
622 are lower.

623 Since the endemic component accounted for such a small minority of the cases, group importations of
624 cases in a given department were rarely observed in the simulations. However, tourism, and local
625 events lead to large gatherings and can increase the risks of group importations in a limited period of
626 time [36,37]. We simulated the spatial spread following repeated importations in a given department,
627 and highlighted that although large outbreaks in the department of importations were common,

628 nation-wide transmission following these importations was very rare. Only the departments where all
629 cases had been imported, and its neighbours, were at risk of uncontrolled outbreaks. Decreasing the
630 level of vaccination by three percent was associated with a large increase in the level of exposure of all
631 departments, and in the number of departments where large outbreaks were generated (Supplement
632 Section 6 and 7). The high levels of transmission observed in recent years in France suggest that
633 importations are frequent, and even a small drop in vaccination could dramatically increase measles
634 transmission in the country.

635 Furthermore, since the number of inhabitants was strongly associated with risks of background
636 importations, most of the endemic importations were reported in urban areas, where the risks of
637 exportations were lower. This could explain the discrepancies between the distribution of the number
638 of cases in the simulations (Figure 5, top row), and the actual number of cases reported in France in
639 2019 [42]. Active transmission was reported in a number of rural areas, notably in the South West of
640 France, and in Savoie (East). This could be due to importations and cross-regional transmission that are
641 under-estimated by our model. Although the model captured the dynamics seen in the data, the
642 calibration study showed it was only able to predict short-term transmission up to one week. The PIT
643 histogram associated to the 14-day calibration displayed signs of bias, which shows that the model was
644 not able to consistently predict variations in the future number of cases in the next two weeks. We
645 identify several factors that could explain the discrepancies observed for longer term predictions: i) the
646 indicator of local immunity we used was flawed: two-dose coverage would be a better indicator of the
647 proportion of the population that is protected; ii) The sub-regional heterogeneity in coverage and past
648 incidence within the department that could be concealed by NUTS3 aggregated data: because of social
649 groups that rarely mix with one another, or large NUTS regions, large outbreaks in a given community
650 would not be a good indicator of the overall level of immunity in a region. Nevertheless, we believe that
651 the results obtained using limited publicly available covariates are encouraging and we intend to apply
652 this method using more complete data.

653 We identified a number of limitations of this study that have not yet been mentioned: Firstly, potential
654 reactive control measures in case of high transmission were not accounted for. It is likely that if the
655 level of incidence was increasing over a short period of time, control measures would be implemented
656 and the behaviour of the individuals may change (e.g. school closures, catch-up vaccination campaigns).
657 This could impact the number of expected cases after a certain threshold is passed, and impact the
658 dynamics in the simulated outbreaks. Secondly, we did not include information on the age or genotype
659 of the cases. Therefore, unrelated importations in successive time-steps in a given region may be
660 considered as linked by our model, whereas they should be separated. Further development of this
661 method could focus on taking this aspect into account, in order to give information on the number of

662 independent concurrent chains. Thirdly, since this is not a transmission model, some extreme values
663 could trigger unlikely behaviour. For instance, if the vaccination rate would be 100%, we would still
664 expect sporadic transmission. Although this would not be entirely implausible given that only the
665 vaccination coverage in the past three years was taken into account in the models (i.e. even if it was
666 100% coverage, there could be susceptible individuals in different age groups). Finally, the impact of
667 the different covariates on the number of cases was constant through time. For instance, the impact of
668 seasonality may depend on factors such as the weather which may vary each year, which would not be
669 accounted for in the model we developed.

670 We used variables collected in a wide range of settings (regional vaccine coverage, incidence, number
671 of inhabitants, surface), therefore this analysis can be reproduced in other countries to analyse the
672 potential for local transmission as well as the impact of recent incidence and vaccine-induced immunity.
673 Since the case counts data are not publicly available, we share the code used to generate the analysis
674 applied to a simulated dataset on a Github repository: ([https://github.com/alxsrobert/measles-
675 regional-transmission](https://github.com/alxsrobert/measles-regional-transmission)).

676 Data availability

677 The daily case counts data came from the European Surveillance System – TESSy, provided by Santé
678 Publique France and released by ECDC. The data cannot be shared publicly. To make this study as
679 reproducible as possible, we generated simulated case counts in France over the same timespan as the
680 main analysis. The code used to generate the simulated dataset, and all the figures presented in the
681 paper is shared in a Github repository (<https://github.com/alxsrobert/measles-regional-transmission>).
682 This repository also contains the publicly available covariates used in the model (local vaccine coverage,
683 number of inhabitants, surface, distance between departments).

684 Funding

685 AR was supported by the Medical Research Council (MR/N013638/1). SF was supported by a Wellcome
686 Trust Senior Research Fellowship in Basic Biomedical Science (210758/Z/18/Z). AJK was supported by a
687 Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (206250/Z/17/Z).

688 Disclaimer

689 The views and opinions of the authors expressed herein do not necessarily state or reflect those of
690 ECDC. The accuracy of the authors' statistical analysis and the findings they report are not the
691 responsibility of ECDC. ECDC is not responsible for conclusions or opinions drawn from the data
692 provided. ECDC is not responsible for the correctness of the data and for data management, data

693 merging and data collation after provision of the data. ECDC shall not be held liable for improper or
694 incorrect use of the data.

695 Acknowledgments

696 We acknowledge Helen Johnson and Nick Bundle from ECDC, who participated in the analysis plan. We
697 also acknowledge the ECDC and Santé Publique France for collecting and providing the case count data
698 used in this study.

699 Author Contributions

700 AR, SF and AJK developed the method and the analysis plan. AR implemented the analysis, wrote the
701 code and ran the model. AR interpreted the results, with contributions from SF and AJK. AR wrote the
702 first draft and the supplementary material. AR, SF, AJK contributed to the manuscript, all authors
703 approved the final version.

704 References

- 705 [1] Gastañaduy PA, Budd J, Fisher N, Redd SB, Fletcher J, Miller J, et al. A Measles Outbreak in an
706 Underimmunized Amish Community in Ohio. *N Engl J Med* 2016;375:1343–54.
707 doi:10.1056/NEJMoa1602295.
- 708 [2] Woudenberg T, Van Binnendijk RS, Sanders EAM, Wallinga J, De Melker HE, Ruijs WLM, et al.
709 Large measles epidemic in the Netherlands, May 2013 to March 2014: Changing epidemiology.
710 *Eurosurveillance* 2017;22:1–9. doi:10.2807/1560-7917.ES.2017.22.3.30443.
- 711 [3] Funk S, Knapp JK, Lebo E, Reef SE, Dabbagh AJ, Kretsinger K, et al. Combining serological and
712 contact data to derive target immunity levels for achieving and maintaining measles elimination.
713 *BMC Med* 2019. doi:10.1186/s12916-019-1413-7.
- 714 [4] Glasser JW, Feng Z, Omer SB, Smith PJ, Rodewald LE. The effect of heterogeneity in uptake of
715 the measles, mumps, and rubella vaccine on the potential for outbreaks of measles: A modelling
716 study. *Lancet Infect Dis* 2016;16:599–605. doi:10.1016/S1473-3099(16)00004-9.
- 717 [5] Keenan A, Ghebrehewet S, Vivancos R, Seddon D, MacPherson P, Hungerford D. Measles
718 outbreaks in the UK, is it when and where, rather than if? A database cohort study of childhood
719 population susceptibility in Liverpool, UK. *BMJ Open* 2017;7. doi:10.1136/bmjopen-2016-
720 014106.
- 721 [6] World Health Organization (WHO). Global Vaccine Action Plan Global Vaccine Action Plan. Who

- 722 2011:4–7.
- 723 [7] World Health Organization. Framework for verifying elimination of measles and rubella. *Wkly*
724 *Epidemiol Rec* 2013;88:89–100. doi:10.1371/jour.
- 725 [8] World Health Organization (WHO). European Region loses ground in effort to eliminate measles
726 2019.
- 727 [9] Pan American Health Organization. Epidemiological Update: Measles. *Paho/ Who* 2019;2020:1–
728 12.
- 729 [10] Fraser B. Measles outbreak in the Americas. *Lancet (London, England)* 2018;392:373.
730 doi:10.1016/S0140-6736(18)31727-6.
- 731 [11] Litvoc MN, Lopes MIBF. From the measles-free status to the current outbreak in Brasil. *Rev Assoc*
732 *Med Bras* 2019;65:1229–30. doi:10.1590/1806-9282.65.10.1129.
- 733 [12] Dimala CA, Kadia BM, Nji MAM, Bechem NN. Factors associated with measles resurgence in the
734 United States in the post-elimination era. *Sci Rep* 2021;11:1–10. doi:10.1038/s41598-020-
735 80214-3.
- 736 [13] Bernadou A, Astrugue C, Méchain M, Le Galliard V, Verdun-Esquer C, Dupuy F, et al. Measles
737 outbreak linked to insufficient vaccination coverage in Nouvelle-Aquitaine region, France,
738 October 2017 to July 2018. *Eurosurveillance* 2018;23:1–5. doi:10.2807/1560-
739 7917.ES.2018.23.30.1800373.
- 740 [14] Held L, Höhle M, Hofmann M. A statistical framework for the analysis of multivariate infectious
741 disease surveillance counts. *Stat Modelling* 2005;5:187–99. doi:10.1191/1471082X05st098oa.
- 742 [15] Meyer S, Held L, Höhle M. hhh4: Endemic-epidemic modeling of areal count time series. *J Stat*
743 *Softw* 2016.
- 744 [16] Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual
745 variation on disease emergence. *Nature* 2005;438:355–9. doi:10.1038/nature04153.
- 746 [17] Fine PEM. The Interval between Successive Cases of an Infectious Disease. *Am J Epidemiol*
747 2003;158:1039–47. doi:10.1093/aje/kwg251.
- 748 [18] Bjørnstad ON, Finkenstädt BF, Grenfell BT. Dynamics of measles epidemics: Estimating scaling
749 of transmission rates using a Time series SIR model. *Ecol Monogr* 2002;72:169–84.
750 doi:10.1890/0012-9615(2002)072[0169:DOMEES]2.0.CO;2.

- 751 [19] Bracher J, Held L. Endemic-epidemic models with discrete-time serial interval distributions for
752 infectious disease prediction. *Int J Forecast* 2020. doi:10.1016/j.ijforecast.2020.07.002.
- 753 [20] Woudenberg T, Woonink F, Kerkhof J, Cox K, Ruijs WLM. The tip of the iceberg : incompleteness
754 of measles reporting during a large outbreak in The Netherlands in 2013 – 2014. *Epidemiol Infect*
755 2018;146:716–22. doi:<https://doi.org/10.1017/S0950268818002698>.
- 756 [21] Lenormand M, Bassolas A, Ramasco JJ. Systematic comparison of trip distribution laws and
757 models. *J Transp Geogr* 2016;51:158–69. doi:10.1016/j.jtrangeo.2015.12.008.
- 758 [22] Institut National de la Statistique et des Etudes Economiques. Estimation de la population au 1^{er}
759 janvier 2020 2020. <https://www.insee.fr/fr/statistiques/1893198#consulter> (accessed
760 September 7, 2020).
- 761 [23] Eurostat. European population grid cells 2011.
762 <https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/grids> (accessed September
763 12, 2020).
- 764 [24] Hijmans RJ, Etten J van, Mattiuzzi M, Sumner M, Greenberg JA, Lamigueiro OP, et al. Package
765 “raster.” R 2014.
- 766 [25] Herzog SA, Paul M, Held L. Heterogeneity in vaccination coverage explains the size and
767 occurrence of measles epidemics in German surveillance data. *Epidemiol Infect* 2011;139:505–
768 15. doi:10.1017/S0950268810001664.
- 769 [26] Santé Publique France. Données départementales 2007-2012 de couverture vaccinale rougeole,
770 rubéole, oreillons à 24 mois 2019. [https://www.santepubliquefrance.fr/determinants-de-
771 sante/vaccination/articles/donnees-departementales-2007-2012-de-couverture-vaccinale-
772 rougeole-rubeole-oreillons-a-24-mois](https://www.santepubliquefrance.fr/determinants-de-sante/vaccination/articles/donnees-departementales-2007-2012-de-couverture-vaccinale-rougeole-rubeole-oreillons-a-24-mois) (accessed September 7, 2020).
- 773 [27] Santé Publique France. Estimations des couvertures vaccinales à 24 mois à partir des certificats
774 de santé du 24^e mois, 2004-2007 2010. [https://www.santepubliquefrance.fr/determinants-de-
775 sante/vaccination/articles/donnees-departementales-2013-2017-de-couverture-vaccinale-
776 rougeole-rubeole-oreillons-a-24-mois](https://www.santepubliquefrance.fr/determinants-de-sante/vaccination/articles/donnees-departementales-2013-2017-de-couverture-vaccinale-rougeole-rubeole-oreillons-a-24-mois) (accessed September 7, 2020).
- 777 [28] Santé Publique France. Données départementales 2013-2017 de couverture vaccinale rougeole,
778 rubéole, oreillons à 24 mois 2019. [https://www.santepubliquefrance.fr/determinants-de-
779 sante/vaccination/articles/donnees-departementales-2013-2017-de-couverture-vaccinale-
780 rougeole-rubeole-oreillons-a-24-mois](https://www.santepubliquefrance.fr/determinants-de-sante/vaccination/articles/donnees-departementales-2013-2017-de-couverture-vaccinale-rougeole-rubeole-oreillons-a-24-mois) (accessed September 7, 2020).
- 781 [29] Antona D, Lévy-Bruhl D, Baudon C, Freymuth F, Lamy M, Maine C, et al. Measles elimination

- 782 efforts and 2008-2011 outbreak, France. *Emerg Infect Dis* 2013;19:357–64.
783 doi:10.3201/eid1903.121360.
- 784 [30] Institut de Veille Sanitaire. Données de déclaration obligatoire de la rougeole. 2009.
- 785 [31] Fitzpatrick G, Ward M, Ennis O, Johnson H, Cotter S, Carr MJ, et al. Use of a geographic
786 information system to map cases of measles in real-time during an outbreak in Dublin, Ireland,
787 2011. *Eurosurveillance* 2012;17:1–11. doi:10.2807/ese.17.49.20330-en.
- 788 [32] Yang W, Wen L, Li SL, Chen K, Zhang WY, Shaman J. Geospatial characteristics of measles
789 transmission in China during 2005–2014. *PLoS Comput Biol* 2017;13:1–21.
790 doi:10.1371/journal.pcbi.1005474.
- 791 [33] Andrianou XD, Del Manso M, Bella A, Vescio MF, Baggieri M, Rota MC, et al. Spatiotemporal
792 distribution and determinants of measles incidence during a large outbreak, Italy, september
793 2016 to july 2018. *Eurosurveillance* 2019;24:1–12. doi:10.2807/1560-
794 7917.ES.2019.24.17.1800679.
- 795 [34] Funk S, Camacho A, Kucharski AJ, Lowe R, Eggo RM, Edmunds WJ. Assessing the performance of
796 real-time epidemic forecasts: A case study of Ebola in the Western Area Region of Sierra Leone,
797 2014–15. *BioRxiv* 2017:1–17. doi:10.1101/177451.
- 798 [35] Czado C, Gneiting T, Held L. Predictive Model Assessment for Count Data 2009:1254–61.
799 doi:10.1111/j.1541-0420.2009.01191.x.
- 800 [36] le Polain de Waroux O, Saliba V, Cottrell S, Young N, Perry M, Bukasa A, et al. Summer music and
801 arts festivals as hot spots for measles transmission: Experience from England and Wales, June
802 to October 2016. *Eurosurveillance* 2016;21:1–6. doi:10.2807/1560-7917.ES.2016.21.44.30390.
- 803 [37] Gautret P, Steffen R. Communicable diseases as health risks at mass gatherings other than Hajj:
804 What is the evidence? *Int J Infect Dis* 2016;47:46–52. doi:10.1016/j.ijid.2016.03.007.
- 805 [38] Patel MK, Goodson JL, Alexander JP, Kretsinger K, Sodha S V, Steulet C. Progress Toward Regional
806 Measles Elimination — Worldwide , 2000 – 2019 2020;69:1700–5.
- 807 [39] Blumberg S, Enanoria WTA, Lloyd-Smith JO, Lietman TM, Porco TC. Identifying postelimination
808 trends for the introduction and transmissibility of measles in the United States. *Am J Epidemiol*
809 2014;179:1375–82. doi:10.1093/aje/kwu068.
- 810 [40] Saxena S, Skirrow H, Bedford H. Routine vaccination during covid-19 pandemic response. *BMJ*
811 2020;369. doi:10.1136/bmj.m268.

812 [41] Dinleyici EC, Borrow R, Safadi MAP, van Damme P, Munoz FM. Vaccines and routine
813 immunization strategies during the COVID-19 pandemic. Hum Vaccines Immunother
814 2021;17:400–7. doi:10.1080/21645515.2020.1804776.

815 [42] Santé Publique France. Bulletin épidémiologique rougeole. Données de surveillance 2019. 2020.
816 <https://www.santepubliquefrance.fr/content/download/231366/2508985> (accessed May 3,
817 2021).

818