Population modeling of early COVID-19 epidemic dynamics in French regions and estimation of the lockdown impact on infection rate

Mélanie Prague[1,2], Linda Wittkop[1,2,3], Quentin Clairon[1,2], Dan Dutartre[4], Rodolphe Thiébaut[1,2,3,†], and Boris P. Hejblum[1,2]

[1]University of Bordeaux, Inria Bordeaux Sud-Ouest, Inserm, Bordeaux Population Health Research Center, SISTM Team, UMR 1219, F-33000 Bordeaux, France
[2]Vaccine Research Institute, F-94000 Créteil, France
[3]CHU Bordeaux, F-33000 Bordeaux, France
[4]Inria Bordeaux Sud-Ouest, F-33000 Bordeaux, France
[†]rodolphe.thiebaut@u-bordeaux.fr

Abstract

We propose a population approach to model the beginning of the French COVID-19 epidemic at the regional level. We rely on an extended Susceptible-Exposed-Infectious-Recovered (SEIR) mechanistic model, a simplified representation of the average epidemic process. Combining several French public datasets on the early dynamics of the epidemic, we estimate region-specific key parameters conditionally on this mechanistic model through Stochastic Approximation Expectation Maximization (SAEM) optimization using `Monolix` software. We thus estimate basic reproductive numbers by region before isolation (between 2.4 and 3.1), the percentage of infected people over time (between 2.0 and 5.9% as of May 11th, 2020) and the impact of nationwide lockdown on the infection rate (decreasing the transmission rate by 72% toward a R_e ranging from 0.7 to 0.9). We conclude that a lifting of the lockdown should be accompanied by further interventions to avoid an epidemic rebound.

Keywords: Compartment model; Coronavirus Disease 2019; COVID-19; France; Population modeling; Non-pharmaceutical intervention

1 Introduction

In December 2019, grouped pneumonia cases have been described in the Hubei province, China and SARS-CoV2 was identified on January, 7th as the cause of this outbreak (Li et al., 2020a; Zhu et al., 2020). SARS-CoV2 causes the viral disease which has been named COVID-19 (World Health Organization, 2020b).

SARS-CoV2 rapidly spread all over the world and the pandemic stage was declared on March 11th by the World Health Organization (2020c). On April 28th, over 1,773,084 cases (in accordance with the applied case definitions and testing strategies in the affected countries) including 111,652 deaths have been reported (World Health Organization, 2020a).

The first case in France was declared on January, 24th (Bernard-Stoecklin et al., 2020) and on April 13th, *Santé Publique France* reported 98,076 confirmed cases and 14,967 hospital deaths due to COVID-19.

COVID-19 includes non-specific symptoms such as fever, cough, headache, and specific symptoms such as loss of smell and taste (Gane et al., 2020; Greenhalgh et al., 2020). The virus is transmitted through droplets and close unprotected contact with infected cases. The majority (around 80 %) of infected cases have a mild form (upper respiratory infection symptoms) without specific needs in terms of care. Around 20 % of cases need hospitalization and among those are severe forms (severe respiratory distress) which will need to be admitted to intensive care units (ICU) with potential need of mechanical ventilation. The percentage of patients in need for ICU care varies between 5 % reported from China (Guan et al., 2020) and 16 % reported from Italy (Grasselli et al., 2020). The number of ICU beds in France was 5,058 at the end of 2018 (DREES, 2019) (although it is currently being increased, having doubled and aiming to reach 14,000 according to the French minister of Health). Thus, the availability of ICU beds with mechanical ventilation is one of the major issues as facilities are not prepared to deal with the potential increase of the number of patients due to this epidemic.

Unprecedented public-health interventions have been taken all over the world (Kraemer et al., 2020) to tackle this epidemic. In France, interventions such as heightening surveillance with rapid identification of cases, isolation, contact tracing, and follow-up of potential contacts were initially implemented. But as the epidemic continued growing, comprehensive physical distancing measures have been applied since March 15th, 2020 including closing of restaurants, non-vital business, schools and universities *etc*, quickly followed by state-wide lockdown on March 17th 2020. The president has announced on April 13th 2020, a progressive lifting of the lockdown from May

11th 2020 onwards. In Wuhan (Hubei, China), the extremely comprehensive physical distancing measures in place since January 23rd have started to be relaxed after 2 months of quarantine and lifted completely on April 8th 2020 (Tian et al., 2020; Wu and McGoogan, 2020).

Interestingly, these interventions have been informed by mathematical models used to estimate the epidemic key parameters as well as unmeasured compartments such as the number of infected people. Another interesting outcome is the forecast of the COVID-19 epidemic according to potential interventions.

Several models have already been proposed to model and forecast the COVID-19 epidemic using compartment models (Fang et al., 2020; Tang et al., 2020; Wang et al., 2020) or agent based models (Di Domenico et al., 2020a; Ferguson et al., 2020; Wilder et al., 2020), its potential impact on intensive care systems (Fox et al., 2020; Massonnaud et al., 2020), and to estimate the effect of containment measurements on the dynamics of the epidemic (Magal and Webb, 2020; Prem et al., 2020). Most of those rely on simulations with fixed parameters and do not perform direct statistical estimations from incident data (Massonnaud et al., 2020). Roques et al. (2020) used French national data but did not use a population approach to model the epidemic at a finer geographical granularity. Yet, the dynamics of the epidemic can be very heterogeneous between regions inside a given country resulting in tremendous differences in terms of needs for hospital and ICU beds (Massonnaud et al., 2020). Moreover, the data collection yields noisy observations, that we deal with a statistical modeling of the observation process, rather than altering the data by e.g. smoothing such as in Roques et al. (2020).

In the present study, we use public data from the COVID-19 outbreak in France to estimate the dynamics of the COVID-19 epidemic in France at the regional level. We model the epidemic with a SEIRAH model, which is an extended Susceptible-Exposed-Infectious-Recovered ($SEIR$) model accounting for time-varying population movements, non-reported infectious subjects (A for unascertained) and hospitalized subjects (H) as proposed by Wang et al. (2020) to model the epidemic in Wuhan. Parameters from the model are estimated at the regional scale using a population approach which allows for borrowing information across regions, increasing the amount of data and thereby strengthening the inference while allowing for local disparities in the epidemic dynamics. Furthermore, we use forward simulations to predict the effect of non-pharmaceutical interventions (NPI) (such as lift of lockdown) on ICU bed availability and on the evolution of the epidemic. Section 2 introduces the data, the model and the necessary statistical tools, Section 3

presents our results and Section 4 discusses our findings and their limits.

2 Methods

Because epidemics spread through direct contacts, their dynamics have a strong spatial component. While traditional compartment models do not account for spatiality, we propose to take it into account by: i) modeling the epidemic at a finer, more homogeneous geographical scale (this is particularly important once lockdown is in place); ii) by using a population approach with random effects across French regions which allows each region to have relatively different dynamics while taking all information into account for the estimation of model parameters iii) aligning the initial starting time of the epidemic for all regions. The starting date in each region was defined as the first date with incident confirmed cases of COVID-19 directly followed by 3 additional consecutive days with incident confirmed cases as well. This criterion of 4 consecutive days with incident cases is needed in particular for the *Île-de-France* region which had 3 consecutive days with 1 imported confirmed case in late January which did not lead to a spreading outbreak at that time.

2.1 Data sources

Open-data regarding the French COVID-19 epidemic is currently scarce, as the epidemic is still unfolding. *Santé Publique France* (SPF) in coordination with the French regional health agencies (*Agences Régionales de Santé* – ARS) has been reporting a number of aggregated statistics at various geographical resolutions since the beginning of the epidemic. During the first weeks of the epidemic in France, SPF was reporting the cumulative number of confirmed COVID-19 cases with a positive PCR test. Other French surveillance resources such as the *Réseau Sentinelles* (Valleron et al., 1986) or the SurSaUD® database (Caserio-Schönemann et al., 2014) quickly shifted their focus towards COVID-19, leveraging existing tools to monitor the ongoing epidemic in real time, making as much data available as possible (given privacy concerns). In this study, we combined data from three different open-data sources: i) the daily release from SPF; ii) the SurSaUD® database that started recording visits to the Emergency room for suspicion of COVID-19 on February, 24th; iii) the *Réseau Sentinelles* which started estimating the weekly incidence of COVID-19 in each French region on March 16th. From the daily release of SPF, we computed the daily incident number of confirmed COVID-19 cases (i.e. with a positive PCR test) in each region. In

addition we used the incident number of visits to the emergency room for suspicion of COVID-19 in each region from the SurSaUD® database using the OSCOUR® network that encompasses more than 86% of all French emergency services (Caserio-Schönemann et al., 2014). Although this does not represent the full extend of hospitalized COVID-19 cases, it is the only public data available that early in the epidemic, when the majority of COVID-19 cases at hospitals were admitted through emergency rooms. Finally, we used the *Réseau Sentinelles* network's weekly incidence estimates of symptomatic cases (including non confirmed cases) to set the ratio between ascertained and unascertained cases in each region (later denoted as r_i). Table 1 presents these observed data. Of note, we studied the epidemic in the 12 Metropolitan French regions – excluding the Corsican region (*Corse*) which exhibits different epidemic dynamics, possibly due to its insular nature.

2.2 Model

2.2.1 Structural model of the epidemic

Wang et al. (2020) extended the classic SEIR model to differentiate between different statuses for infected individuals: ascertained cases, unascertained cases and cases quarantined by hospitalization. The model, assuming no population movement, is presented in Figure 1. The population is divided into 6 compartments: susceptible S, latent E, ascertained infectious I, unascertained infectious A, hospitalized infectious H, and removed R (recovered and deceased). This model assumes that infections are well-mixed throughout the population, ignoring any spatial structure or compartmentalization by population descriptors such as age. Such assumptions make it particularly relevant to infer the dynamics of the French epidemic at the regional level (a finer geographical scale at which such hypotheses are more likely to hold). Figure 1 illustrates the dynamics between those 6 compartments that are characterized by the following system of six Ordinary Differential Equations (ODE):

$$
\begin{cases}
\dfrac{dS}{dt} &= -\dfrac{bS(I+\alpha A)}{N} \\
\dfrac{dE}{dt} &= \dfrac{bS(I+\alpha A)}{N} - \dfrac{E}{D_e} \\
\dfrac{dI}{dt} &= \dfrac{rE}{D_e} - \dfrac{I}{D_q} - \dfrac{I}{D_I} \\
\dfrac{dR}{dt} &= \dfrac{(I+A)}{D_I} + \dfrac{H}{D_h} \\
\dfrac{dA}{dt} &= \dfrac{(1-r)E}{D_e} - \dfrac{A}{D_I} \\
\dfrac{dH}{dt} &= \dfrac{I}{D_q} - \dfrac{H}{D_h}
\end{cases}
\tag{1}
$$

Model parameters are described in Table 2. Of note, given a combination of parameters and initial states of the system $\xi = (b, D_q, r, \alpha, D_e, D_I, D_h, N, S(t=0),\ E(t=0), I(t=0), R(t=0), A(t=0), H(t=0))$, using a solver of differential equations, it is possible to deterministically compute at any time t the quantities $S(t,\xi)$, $E(t,\xi)$, $I(t,\xi)$, $R(t,\xi)$, $A(t,\xi)$, and $H(t,\xi)$.

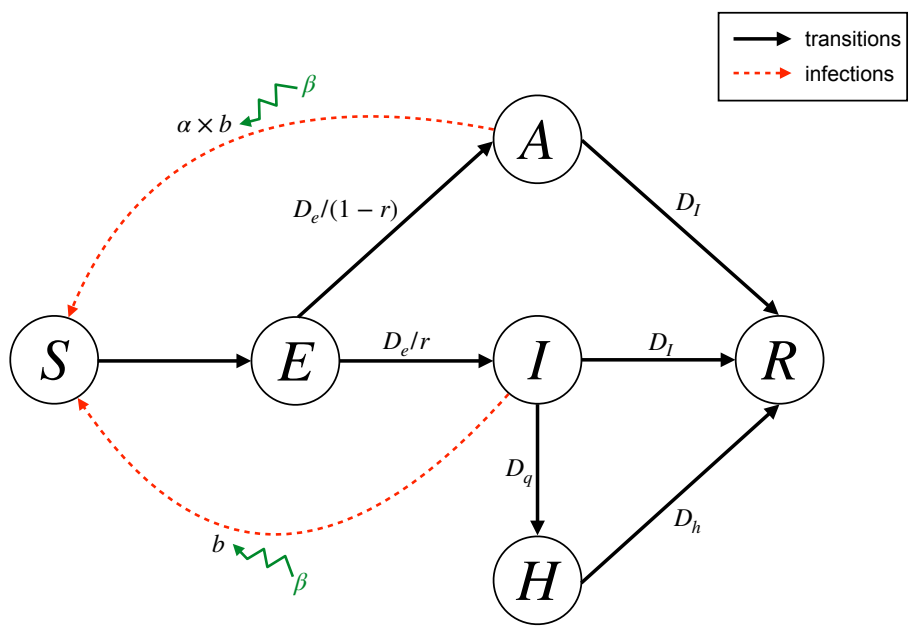

Figure 1: SEIRAH model representation – adapted from Wang et al. (2020)

Geographical region	**Epidemics start date**	**Observed cumulative Ascertained cases until 2020-03-25**	**Observed cumulative Hospitalized cases until 2020-03-25**	**Population Size**	**ICU Capacities on 2018-12-31**	**Percentage of ascertained cases from** *Réseau Sentinelles* r_s
Auvergne-Rhône-Alpes	2020-03-02	2,082	1,456	8,032,377	559	0.03
Bourgogne-Franche-Comté	2020-03-02	1,565	1,095	2,783,039	198	0.07
Bretagne	2020-03-02	602	273	3,340,379	162	0.02
Centre-Val de Loire	2020-03-10	545	138	2,559,073	180	0.03
Grand Est	2020-03-02	5,478	2,435	5,511,747	465	0.05
Hauts-de-France	2020-03-02	1,744	762	5,962,662	438	0.02
Île-de-France	2020-03-02	7,651	3,870	12,278,210	1,147	0.04
Nouvelle Aquitaine	2020-03-04	908	695	5,999,982	412	0.03
Normandie	2020-03-10	677	214	3,303,500	240	0.02
Occitanie	2020-03-02	1,081	673	5,924,858	474	0.03
Provence-Alpes-Côte d'Azur	2020-03-02	1,927	1,241	5,055,651	460	0.05
Pays de la Loire	2020-03-04	363	536	3,801,797	181	0.01
France	NA	24,623	13,388	64,553,275	4,916	0.03 [0.01;0.07]

Table 1: Description of the COVID-19 epidemic by French region. The epidemic start date is defined as the first day followed by three consecutive days with at least 1 confirmed case.

Parameter	Interpretation	Value	References
b	Transmission rate of ascertained cases	region specific	*estimated*
r	Ascertainement rate	region specific	*Réseau Sentinelles*
α	Ratio of transmission between A and I	0.55	Li et al. (2020b)
D_e	Latent (incubation) period (days)	5.1	Lauer et al. (2020)
D_I	Infectious period (days)	2.3	Li et al. (2020a); Wang et al. (2020)
D_q	Duration from I onset to H (days)	region specific	*estimated*
D_h	Hospitalization period (days)	30	Li et al. (2020a); Wang et al. (2020)
N	Population size	region specific	INED (2020)
β	Effect of isolation	-	*estimated*

Table 2: Parameters of the SEIRAH model

2.2.2 Observation model

Observation processes In our case, none of the compartments of the system are directly observed: the only observations considered are i) the number of daily incident infectious ascertained cases denoted Y^1, and ii) the number of daily incident hospitalized infectious cases denoted Y^2. These observations are the only one available both before and after the initiation of lockdown. Those two quantities are modeled in Equation (1) respectively as observations from the $I^{(in)}(t,\xi) = \frac{rE(t,\xi)}{D_e}$ and $H^{(in)}(t,\xi) = \frac{I(t,\xi)}{D_q}$ random variables, which are the numbers of new incident cases at time t given the parameters ξ in compartment I and H respectively. Because these are count processes, we propose to model their observations Y^1 and Y^2 with Poisson likelihoods:

$$\begin{aligned} P(Y_i^1 = k_1) &= \frac{e^{-I^{(in)}(t,\xi)} {I^{(in)}}^{k_1}(t,\xi)}{k_1!} \\ P(Y_i^2 = k_2) &= \frac{e^{-H^{(in)}(t,\xi)} {H^{(in)}}^{k_2}(t,\xi)}{k_2!} \end{aligned} \tag{2}$$

where k_1 and k_2 are the respective numbers of cases.

Initial values The initial states of all compartments at the date of epidemic start ($t = 0$) for region i are also important drivers of the dynamics. Some of them can be approximated by quantities directly depending on the observations: i) $I_i(t = 0) = \sum_{t \leq 0} Y_i^1(t)$, ii) $H_i(t = 0) = \sum_{t \leq 0} Y_i^2(t)$, and iii) $R_i(t = 0) = 0$. Others, namely compartments A and E are not directly observed, and we evaluate these initial quantities. Due to variation in data collection protocols and the initial size of regional outbreaks, this estimation is particularly important. Indeed, the number of daily incident cases at $t = 0$ ranges from 1 to 37 cases depending on the region. $A_i(t = 0)$ is set as $\frac{I_i(t=0)(1-r)}{r}$ by model assumption. $E_i(t = 0)$ is estimated as a direct parameter of the model (see Section 2.2.3). Finally, because the total population size is equal to N_i, we have $S_i(t = 0) = N_i - E_i(t = 0) - I_i(t = 0) - R_i(t = 0) - A_i(t = 0) - H_i(t = 0)$.

2.2.3 Statistical population model

The goal of this study is to model the epidemic of COVID-19 in France, but at the regional level using a population approach. This is done using a mixed effect model. In this inference framework, baseline parameters governing the dynamics of the epidemic in each region are assumed to be drawn from a shared distribution which allows for heterogeneity between regions, known as the random effects. We use the log-normal distribution for all parameters to ensure their positivity during estimation. Because public health policies changed over the time period of observation of the epidemic, we incorporate explanatory covariates such as physical distancing by lockdown (C^1 and C^2) as a time-dependent effect on the transmission of the disease b. Covariate C^1 is 0 until 2020-03-17 date of the start of the policy in France and then set to 1. Covariate C^2 is 0 until 2020-03-25 assuming that social distancing behaviours build up in a week. In other words, we have $\forall i = 1, \ldots, 12$ (where i is the region identifier):

$$\begin{array}{lcll} log(b_i(t)) & = & b_0 + \beta_1 C_i^1(t) + \beta_2 C_i^2(t) + u_i^b, & u_i^b \sim \mathcal{N}(0, \sigma_b^2) \\ log(D_{q_i}) & = & D_{q0} + u_i^{D_q}, & u_i^{D_q} \sim \mathcal{N}(0, \sigma_{D_q}^2) \\ log(E_i(t=0)) & = & E(t=0) + u_i^{E_0}, & u_i^{E_0} \sim \mathcal{N}(0, \sigma_{E_0}^2) \end{array} \quad (3)$$

The parameters $(b_0, D_{q0}, E(t = 0))$ are mean shared values in the population, and can be seen as the country values for these parameters. The inter-region random-effect $(u_i^b, u_i^{D_q}, u_i^{E_0})$ are normally distributed and assumed independent. So the vector of parameters in the model for each region is $\xi_i = (b_i, D_{q_i}, \beta_1, \beta_2, r_i, \alpha, D_e, D_I, D_h, N_i, S_i(t = 0), E_i(t = 0), I_i(t = 0), R_i(t =$

$0), A_i(t = 0), H_i(t = 0))$. β_1 can be interpreted through $K_1 = exp(-\beta_1)$ which is the factor by which transmission is modified after the start of lockdown during the first week. The factor by which transmission is modified after that first week of confinement is given by $K_2 = exp(-\beta_1 - \beta_2)$. The coefficients $\beta = (\beta_1, \beta_2)$ are expected to be negative as lockdown aims at reducing transmission. Interestingly, with our approach, we can evaluate whether or not there is a statistically significant effect of lockdown on the transmission by testing $\beta = 0$ using a Wald test.

2.3 Inference

2.3.1 Estimation

Region-specific model parameters Based on the results from the theoretical identifiability analysis of the structural model of the epidemic from Equation (1) (see Appendix A2), we estimate the parameters (b_i, D_{q_i}, r_i) as well as the initial state (E_{0i}) when the epidemic begins being reported in each region i. We used the `Monolix` software version 2019R2 (Lixoft SAS, 2019) to estimate those five parameters by maximizing the likelihood of the data given the model and the other fixed parameters. This software relies on a frequentist version of the Stochastic Approximation Expectation Maximization (SAEM) algorithm (Delyon et al., 1999) and standard errors are calculated via estimation of the Fisher Information Matrix (Kuhn and Lavielle, 2005), which is derived from the second derivative of the log-likelihood evaluated by importance sampling. In addition, we use profile-likelihood to confirm that no further information can be gained from the data at hand on parameters α, D_e and D_I by running the SAEM algorithm multiple times while setting these parameters to different values and obtaining similar maximum likelihood values (meaning more data would be needed to be able to estimate those parameters). During inference, practical identifiability of the model is evaluated by the ratio of the minimum and maximum eigenvalues of the Fisher Information Matrix, that will be referred as "convergence ratio" in the reminder of the manuscript. Convergence of the SAEM algorithm was assessed by running multiple SAEM chains and checking that they all mix around similar probability distributions.

Compartments dynamics We are particularly interested in the trajectories of the model compartments. We use Monte Carlo methods (parametric bootstrapping) to compute the confidence intervals accounting for the uncertainty in estimating the structural and statistical model parameters. For

all compartments $C(t, \xi_i)$ (C being S, E, I, R, A, or H) the 95% confidence interval is estimated by sampling from the posterior distributions of the model parameters to simulate 1,000 trajectories, and taking the 2.5% and 97.5% percentiles of these simulated trajectories. We also added to it the error measurement given by the Poisson distribution of the outcomes. Other outcomes of interest are the number of ICU beds needed and the number of death (D) in a given region at a given time. These quantities are not specifically modeled by our mathematical structural model. However, it is possible to roughly approximate them by assuming that they represent a percentage of the hospitalized cases $H(t, \xi_i)$ and removed cases $R(t, \xi_i)$. We assume that $ICU(t, \xi_i) = 0.25 \times H(t, \xi_i)$ which is consistent with the prevalence of ICU cases among hospitalized cases at the French national level. Based on the estimation of the Infection Fatality Ratio (IFR) from Roques et al. (2020), we get a rough estimation of $D(t, \xi_i)$ as 0.5% of $R(t, \xi_i)$. Roques et al. (2020) conclude that COVID-19 fatalities are under-reported, and using their IFR estimate we adequately fit the trend of the observed COVID-19 deaths but with an offset due to this assumed higher IFR, see Appendix A1.

Model update Furthermore β estimations can be easily updated as new data become available using parametric empirical Bayes (thus avoiding the need to re-estimate the whole system). It consists in maximizing the likelihood again with respect to β while holding the other parameter distribution fixed to their previously inferred *a posteriori* distribution. This is how our results are updated with data after March 25th 2020 in this work.

Effective reproductive number For each region, we compute the effective reproductive number $R_e(t, \xi_i)$ as a function of model parameters:

$$R_e(t, \xi_i) = \frac{D_I b_i}{A(t, \xi_i) + I(t, \xi_i)} \left(\alpha A(t, \xi_i) + \frac{D_{q_i} I(t, \xi_i)}{D_I + D_{q_i}} \right) \tag{4}$$

When individuals are homogeneous and mix uniformly, $R_e(t, \xi_i)$ is defined as the mean number of infections generated during the infectious period of a single infectious case in the region i at time t. This is the key parameter targeted by NPIs. We compute analytically its 95% confidence interval by accounting for all 95% confidence interval $[X^{min}; X^{max}]$ of parameters and

trajectories X used in its definition such that:

$$CI_{95\%}(R_e(t,\xi_i)) = \left[\frac{D_I b_i^{min}}{A^{max}(t,\xi_i)+I^{max}(t,\xi_i)}\left(\alpha A^{min}(t,\xi_i)+\frac{D_{q_i}^{min} I^{min}(t,\xi_i)}{D_I+D_{q_i}^{max}}\right);\right. \quad (5)$$
$$\left.\frac{D_I b_i^{max}}{A^{min}(t,\xi_i)+I^{min}(t,\xi_i)}\left(\alpha A^{max}(t,\xi_i)+\frac{D_{q_i}^{max} I^{max}(t,\xi_i)}{D_I+D_{q_i}^{min}}\right)\right]$$

Asymptomatic proportion At a given time t the number of incident unascertained cases is equal to the sum of two populations, the number of incident non-tested symptomatic individuals (NT) and the number of incident non-tested asymptomatic individuals (AS):

$$\frac{(1-r)E(t,\xi)}{D_e} = NT(t,\xi)+AS(t,\xi) \quad (6)$$

where r is the proportion of cases tested positive. Collection of data from general practitioners through the re-purposing of the *Réseau Sentinelles* network to monitor COVID-19 provides a weekly estimation of the number of incident symptomatic cases (tested or not tested) that we previously called r_s. This quantity is given over a week but can be evaluated daily by averaging:

$$r_s = \frac{\frac{rE(t,\xi)}{D_e}}{NT(t,\xi)+\frac{rE(t,\xi)}{D_e}} \quad (7)$$

where r_s represents the proportion of infected cases seeing a general practitioner. Combining equations (6) and (7) allows to compute the incident number of asymptomatic cases as a function of the compartment E:

$$AS(t,\xi) = \frac{E(t,\xi)}{D_e}\frac{1-(r+r_s)}{1-r_s} \quad (8)$$

2.3.2 Predictions

Short-term predictions of attack rates We predict the proportion of infected individuals among the population in each region at a given date by computing: $[E(t,\xi_i)+A(t,\xi_i)\ +I(t,\xi_i)+H(t,\xi_i)+R(t,\xi_i)]/N_i$ (neglecting the deaths).

Targeting lockdown consequences Given the values of parameters ξ_i, we predict the trajectories of the dynamical system compartments using the `dlsode` differential equation solver in `R` (Soetaert et al., 2020) and we can investigate the impact of NPIs such as lockdown in various scenarios. This impact is driven by two parameters:

- K_2 (K_1 is considered fixed), the decrease ratio of transmission rate of the disease following the first week of NPI, defined as $b_i^C = \frac{b_i}{K_2}$. This directly translates into a decrease of the effective reproductive number according to Equation (4). It reflects the fact that individual by getting confined decrease their number of contacts. Of note, the current K_2 is estimated by $exp(-\widehat{\beta}_1 - \widehat{\beta}_2)$ (see Section 2.2.3).

- τ, the duration (in days) of the lockdown during which the mixing and transmission are fixed to $b_i^C = \frac{b_i}{K_2}$ instead of b_i.

We evaluate the magnitude of the possible epidemic rebound after confinement according to several values for K_2. In particular, we predict the rates $E(t, \xi_i)/N_i$, $A(t, \xi_i)/N_i$ and $I(t, \xi_i)/N_i$ on May 11th 2020 (currently considered by French authorities as the possible start date for lifting lockdown in France). We also compute the optimal lockdown duration τ_i^{opt} needed to achieve the epidemic extinction in region i defined as $E_i(t, \xi_i) < 1$ and $A_i(t, \xi_i) < 1$ and $I_i(t, \xi_i) < 1$ simultaneously. In each scenario we predict the date at which the ICU capacities in each region would be overloaded, this date is given by:

$$t_i^{ICU} = \underset{ICU(t,\xi_i)>\eta_i}{\arg\min} \ (t)$$

η_i denoting the ICU capacities limits in region i. We additionally predict how many more ICU beds would be needed at the peak of hospitalization in each scenario, as a proportion of current ICU capacity (DREES, 2019). Finally we provide a rough prediction of the number of deaths for each envisioned scenario assuming a confinement duration of τ_i^{opt} days.

3 Results

3.1 Estimation of the regional epidemic dynamics

Data fitting Because of the lag inherent to diagnostic testing, we also estimated the number of people already infected at this epidemic start by E_0 (notably, the largest numbers of E_0 in Table 3 are estimated for *Île-de-France* and *Grand Est* the two most affected French regions in this early epidemic).

On March 25th, the cumulative number of ascertained cases was 24,623 and the cumulative number of hospitalized cases was 13,388, see Table 1 for a regional breakdown. Our SEIRAH model fits the data well as can be seen in Figure 2. Moreover, the stability of the estimates is good with a convergence ratio of 1.6 (see Section 2.3.1), corroborating the good identifiability of the estimated parameters (see Appendix A2). Table 3 provides the regional estimates of the transmission rates (b_i). Of note, regions with higher transmission rate are not necessarily those known to have the highest number of incident ascertained cases. D_{q_i}, the number of days from illness onset to hospitalization, can be quite variable between regions and likely accounts for heterogeneity in the observed data. We estimates its population mean at $D_{q_0} = 1.13$ days with a standard deviation $\sigma_{D_q} = 0.42$.

To evaluate the validity of our structural ODE model (1) and of our inference results, we compare the aggregated predictions of the number of both incident ascertained cases and incident hospitalized cases at the national French level to the daily observed incidences (that are still publicly available from SPF at the country level, even after March 25th – while incident ascertained cases are not openly availilable at the regional level after March 25th). Figure 3 displays both predictions and observations, illustrating the added value of incorporating data after March 25th as those encompass new information about the epidemic dynamics and characteristics as we approach the peak in most regions (notably *Grand Est* and *Île de France*). Of note, worse fit of the observed hospitalizations by our model can be explained by the data discrepancy: while we use the SurSaUD® data for our inference (which only account for patients arriving through the emergency room), Figure 3 displays the data from *Santé Publique France* which should contain all COVID-19 hospitalizations in France (including hospitalizations not coming from the emergency room, as well as Corsica and the French *Departements d'Outre Mer* that are not taken into account in our model).

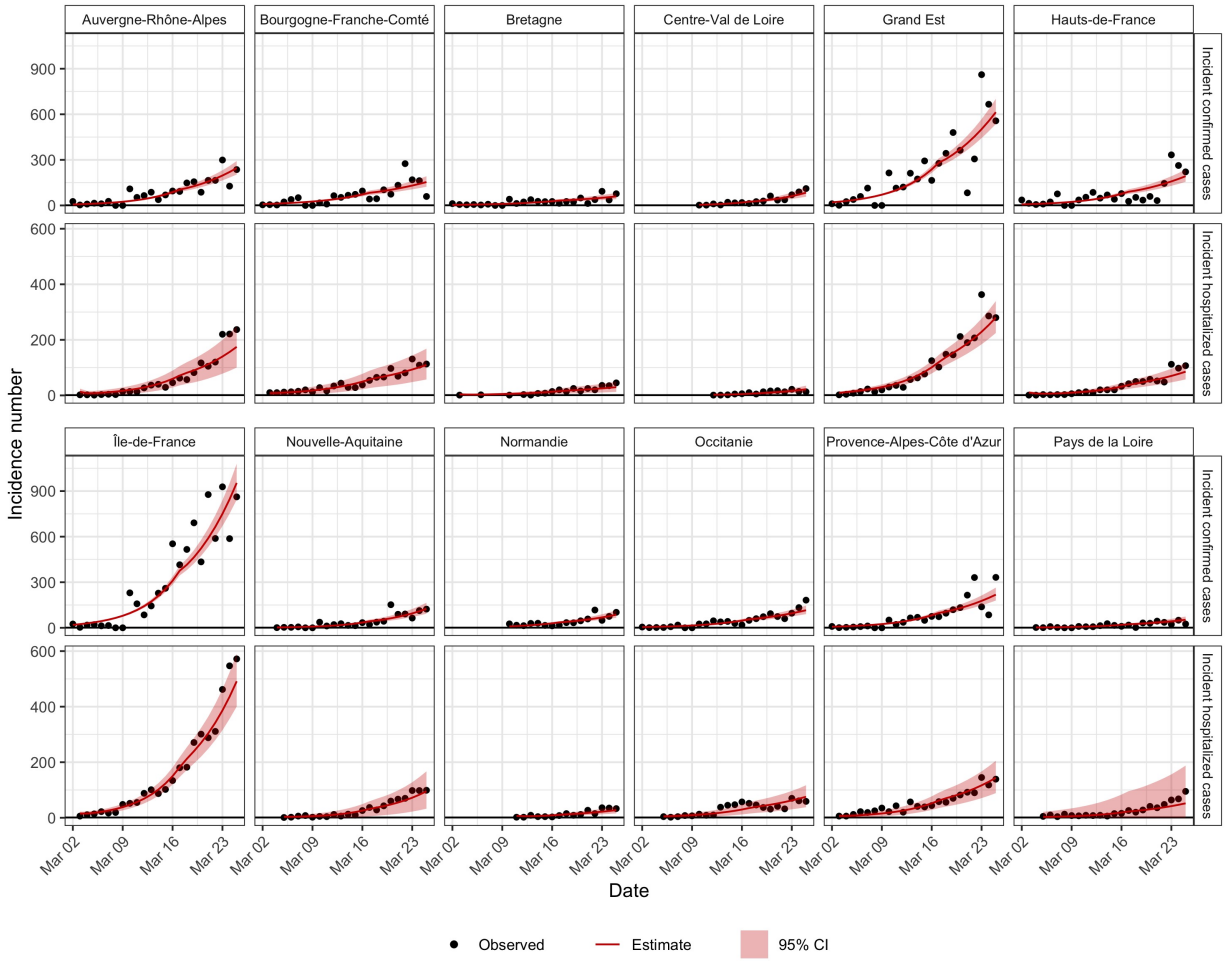

Figure 2: Fitting curves of $I^{(in)}$ and $H^{(in)}$ with the SEIRAH model

Region	**Transmission rate** b_i	E_0**: number of latent cases at epidemic start date**	**Attack rate without confinement** $R(\infty, \xi)/N$
Auvergne-Rhône-Alpes	1.97 [1.94;2.00]	685 [475; 895]	88.7% [87.4%; 89.9%]
Bourgogne-Franche-Comté	1.90 [1.84;1.95]	851 [741; 960]	86.4% [84.6%; 88.1%]
Bretagne	1.71 [1.65;1.78]	855 [567;1,143]	83.6% [81.3%; 85.8%]
Centre-Val de Loire	2.45 [2.34;2.57]	361 [129; 594]	94.8% [93.6%; 95.9%]
Grand Est	2.07 [2.03;2.11]	2,867 [2,577;3,156]	90.6% [89.3%; 91.6%]
Hauts-de-France	1.77 [1.73;1.80]	1,392 [952;1,832]	85.0% [83.3%; 86.5%]
Île-de-France	2.26 [2.22;2.30]	2,414 [2,108;2,720]	92.8% [91.8%; 93.6%]
Nouvelle-Aquitaine	2.46 [2.37;2.56]	388 [274; 503]	94.4% [93.2%; 95.3%]
Normandie	1.90 [1.81;1.98]	2,148 [1,547;2,749]	88.1% [85.9%; 90.0%]
Occitanie	1.99 [1.92;2.06]	1,093 [879;1,307]	89.1% [87.4%; 90.6%]
Provence-Alpes-Côte d'Azur	2.10 [2.04;2.17]	798 [655; 942]	90.4% [89.1%; 91.7%]
Pays de la Loire	2.07 [1.96;2.17]	868 [422;1,314]	90.4% [88.3%; 92.1%]
France	2.10 [2.00;2.10]	1,372 [1,095;1,650]	89.9% [88.5%;91.2%]

Table 3: Estimation of region-wise model parameters as well as the final attack rate (without confinement).

Evolution of the epidemic without intervention It is also interesting to predict the percentage of infected individuals in each region at future dates, corresponding to attack rates. Table 3 provides such attack rates by region in the absence of any NPI, showing that 89.5% [88%; 90.7%] of the French population would end up being infected. In this scenario we estimate that the peak of the epidemic would have been around May 3rd 2020 and that the epidemic would have finally gone extinct around May 25th 2021. At the peak of hospitalization, about 714,259 [360,692; 1,065,226] individuals would have required simultaneous hospitalization.

Proportion of asymptomatic cases With the current model, we have $D_e = 5.1$ and $r = r_s = 3.3\%$ $[1.2\%; 6.6\%]$. Thus we estimate that the percentage of asymptomatic infectious individuals is : 18.1% [16.7%; 19.2%].

3.2 Estimation of the lockdown effect

Change of transmission rate during lockdown The parameter β_1 and β_2 measure the effect of the lockdown before and after a week of adjustment. Both are significantly different from 0 ($p<0.001$) such that the lockdown reduced the transmission rate of COVID-19 by a divisive factor K_1 estimated at 1.31 [1.27; 1.35] during the first week and K_2 estimated at 3.63 [3.48; 3.80] after this first week. Of note, thanks to our update algorithm (see Section 2.3.1), it is possible to update those results rapidly as soon as more data are available to inform which scenario of prediction described in Section 3.3 is the most likely.

Effective reproductive number The above quantities directly impact the effective reproductive number as described in Table 4. The overall effective reproductive number for France is 2.6 [2.4; 2.8] before the lockdown, 2.0 [1.8; 2.1] during the first week of lockdown and 0.7 [0.1;5.3] after March 25th 2020. Figure 4 displays the effective reproductive number trajectories in each region.

3.3 Epidemic dynamics predictions after lockdown lift

In Tables 5 and 6, we vary $K_2 = 3, 5, 10$ (the magnitude of the reduction of transmission during lockdown after the first week). This gradient of simulation is important because the actual French value of K_2 remains currently unknown. In Section 3.2 we showed that we can estimate a lower bound for $K_2 \geq \widehat{K}_2 = 3.48$. From Table 5, we show that the higher K_2 is, the lowest

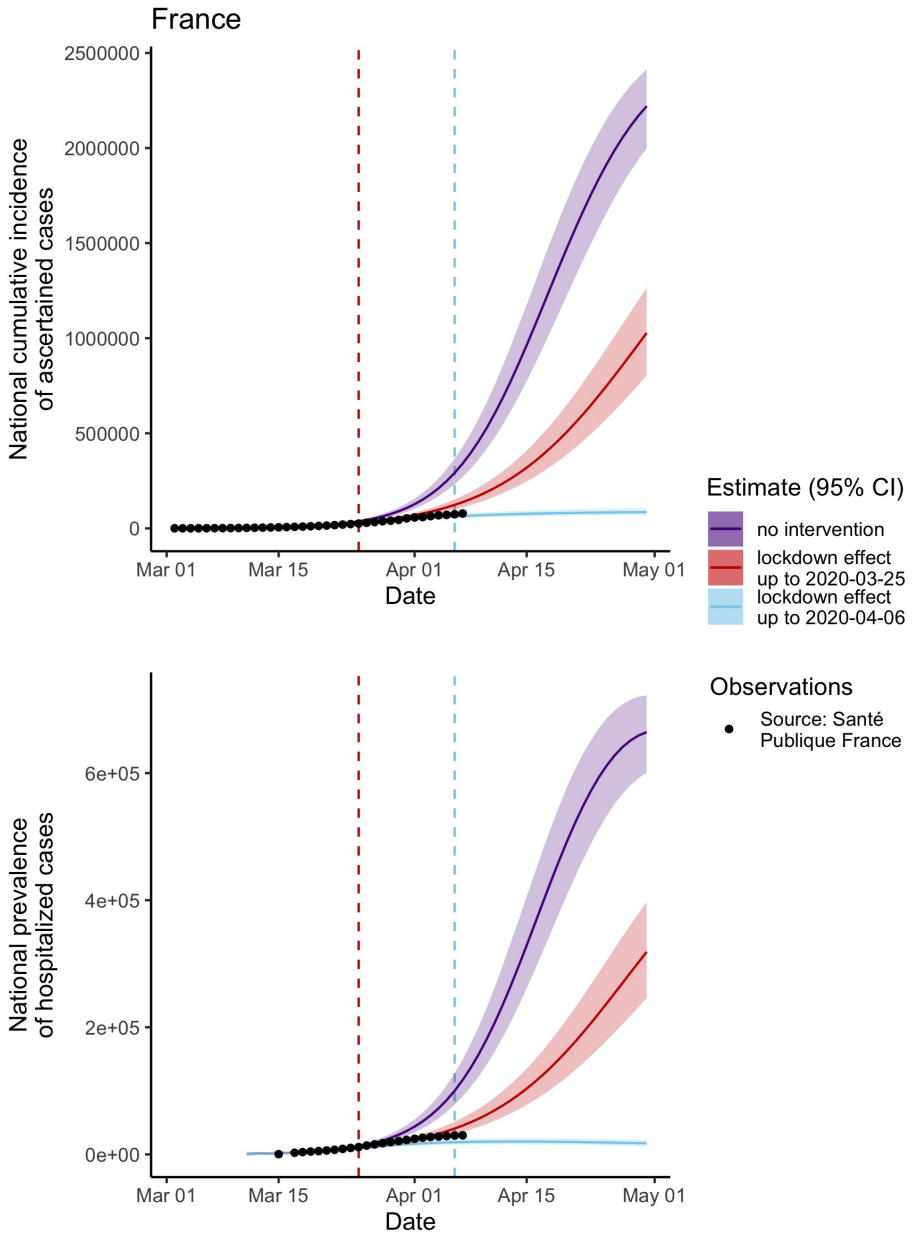

Figure 3: Observed incident number of ascertained cases and hospitalizations in France compared to predicted ones, based on estimations with data collected up to either March 16 (before lockdown), March 25th or April 6th (delimited by the vertical line).

Geographical region	R_e **(before confinement)**	R_e **(after confinement) until 2020-03-25**	R_e **(after confinement) after 2020-03-25**
Auvergne-Rhône-Alpes	2.5 [2.0;3.1]	1.9 [1.5;2.4]	0.7 [0.5;1.0]
Bourgogne-Franche-Comté	2.4 [1.6;3.5]	1.8 [1.2;2.7]	0.7 [0.4;1.1]
Bretagne	2.2 [1.3;3.5]	1.7 [1.0;2.7]	0.6 [0.3;1.2]
Centre-Val de Loire	3.1 [1.9;5.2]	2.4 [1.4;4.0]	0.9 [0.4;1.7]
Grand Est	2.6 [2.0;3.5]	2.0 [1.5;2.7]	0.7 [0.5;1.1]
Hauts-de-France	2.2 [1.7;2.9]	1.7 [1.3;2.3]	0.6 [0.4;0.9]
Île-de-France	2.8 [2.2;3.7]	2.2 [1.7;2.8]	0.8 [0.6;1.1]
Nouvelle-Aquitaine	3.1 [1.8;5.4]	2.4 [1.4;4.1]	0.9 [0.4;1.7]
Normandie	2.4 [1.7;3.5]	1.8 [1.2;2.7]	0.7 [0.4;1.2]
Occitanie	2.5 [1.6;4.0]	1.9 [1.2;3.1]	0.7 [0.4;1.3]
Provence-Alpes-Côte d'Azur	2.6 [1.7;4.1]	2.0 [1.3;3.1]	0.7 [0.4;1.3]
Pays de la Loire	2.6 [1.4;4.8]	2.0 [1.1;3.7]	0.7 [0.3;1.6]
France	2.6 [1.8;3.8]	2.0 [1.4;2.9]	0.7 [0.4;1.2]

Table 4: Estimation of the effective reproductive ratios R_e during each of the 3 considered periods (before lockdown, during the first week of lockdown, and beyond 1 week of lockdown) for each region with 95% confidence intervals.

the numbers of ascertained (I), unascertained (A) and latent (E) infected individuals are on May 11th 2020. However, it is not equal to 0, which means the epidemic is not extinct (and ready to bounce back as soon as lockdown is lifted). In Table 6, we predict the optimal (i.e. shortest) duration of the lockdown to achieve extinction of the epidemic in each region, which is, on average, 407 [236; 786] days if $K_2 = 3$, 147 [123; 179] days if $K_2 = 5$, 97 [88; 112] days if $K_2 = 10$. We now focus on the average scenario in which $K_2 = 5$: only three regions in which ICU capacities will be exceeded (*Île-de-France*, *Grand Est* and *Bourgogne-Franche-Comté*) by the end of March 2020. At these times, ICU capacities will have to be increased by respectively 125% [107%; 146%], 178% [147%; 214%] and 158% [125%; 195%]. In this scenario and lockdown is maintained until τ_i^{opt} in each region, the predicted number of deaths is 86,387 [71,838; 103,969].

Attack rates under lockdown Table 7 presents the proportion of infected individuals at various dates, i.e. instantaneous attack rates. We also predict them for three horizon dates assuming confinement would be maintained until these dates: 2020-05-15, 2020-06-08 and 2020-06-22. We predict the national French attack rate on May 15th 2020 to be 3.8% [3.1%; 4.8%].

Region	Nb. Ascertained Cases (I) on 2020-05-11	Nb. Nonascertained Cases (A) on 2020-05-11	Nb. Latent Cases (E) on 2020-05-11
$K_2 = 3$			
Auvergne-Rhône-Alpes	38 [14; 70]	5,418 [3,651; 7,813]	11,479 [7,719; 16,720]
Bourgogne-Franche-Comté	17 [2; 37]	975 [600; 1,552]	2,131 [1,327; 3,395]
Bretagne	8 [0; 19]	1,035 [595; 1,696]	2,104 [1,224; 3,461]
Centre-Val de Loire	122 [46; 227]	6,191 [3,343; 10,444]	14,030 [7,549; 23,902]
Grand Est	209 [118; 330]	8,187 [5,573; 11,927]	17,889 [12,103; 26,194]
Hauts-de-France	32 [12; 58]	3,472 [2,386; 5,105]	7,074 [4,856; 10,499]
Île-de-France	457 [265; 688]	24,578 [16,739; 34,571]	54,612 [36,837; 77,598]
Nouvelle-Aquitaine	50 [11; 107]	7,982 [4,611; 13,758]	18,185 [10,452; 31,884]
Normandie	28 [7; 57]	2,163 [1,244; 3,605]	4 ,530 [2,608; 7,651]
Occitanie	23 [4; 46]	3,006 [1,732; 4,642]	6,415 [3,704; 9,972]
Provence-Alpes-Côte d'Azur	53 [19; 100]	3,333 [2,073; 5,284]	7,402 [4,632; 11,810]
Pays de la Loire	1 [0; 3]	4,265 [2,345; 7,602]	8,982 [4,946; 16,074]
$K_2 = 5$			
Auvergne-Rhône-Alpes	4 [0; 9]	575 [389; 798]	1,066 [731; 1,473]
Bourgogne-Franche-Comté	2 [0; 5]	110 [58; 174]	210 [120; 326]
Bretagne	1 [0; 3]	128 [62; 210]	229 [119; 369]
Centre-Val de Loire	9 [0; 21]	482 [246; 812]	952 [499; 1,613]
Grand Est	21 [7; 38]	864 [593; 1,252]	1,659 [1,155; 2,405]
Hauts-de-France	4 [0; 9]	429 [290; 616]	770 [534; 1,098]
Île-de-France	39 [16; 65]	2,293 [1,606; 3,100]	4,467 [3,131; 6,049]
Nouvelle-Aquitaine	3 [0; 9]	609 [314; 980]	1,204 [639; 1,943]
Normandie	3 [0; 8]	236 [126; 377]	433 [241; 688]
Occitanie	2 [0; 6]	306 [170; 492]	571 [328; 912]
Provence-Alpes-Côte d'Azur	5 [0; 11]	320 [184; 498]	621 [372; 955]
Pays de la Loire	0 [0; 1]	435 [213; 748]	802 [406; 1,379]
$K_2 = 10$			
Auvergne-Rhône-Alpes	0 [0; 2]	72 [42; 105]	115 [72; 165]
Bourgogne-Franche-Comté	0 [0; 1]	15 [4; 27]	25 [9; 42]
Bretagne	0 [0; 1]	19 [5; 35]	30 [10; 52]
Centre-Val de Loire	1 [0; 3]	42 [16; 73]	71 [31; 120]
Grand Est	2 [0; 6]	103 [63; 150]	171 [110; 243]
Hauts-de-France	1 [0; 2]	63 [35; 92]	97 [59; 140]
Île-de-France	4 [0; 8]	240 [165; 330]	400 [283; 545]
Nouvelle-Aquitaine	0 [0; 1]	54 [22; 91]	91 [42; 149]
Normandie	0 [0; 2]	30 [12; 52]	48 [22; 80]
Occitanie	0 [0; 1]	37 [14; 64]	60 [26; 99]
Provence-Alpes-Côte d'Azur	0 [0; 2]	37 [16; 62]	61 [30; 99]
Pays de la Loire	0 [0; 0]	51 [19; 93]	81 [34; 145]

Table 5: Epidemic state on May 11th 2020. For each region depending on K_2 the effect of lockdown on transmission, we present the predicted number of ascertained (I), unascertained (A) and latent (E) individuals infected and its 95% confidence intervals.

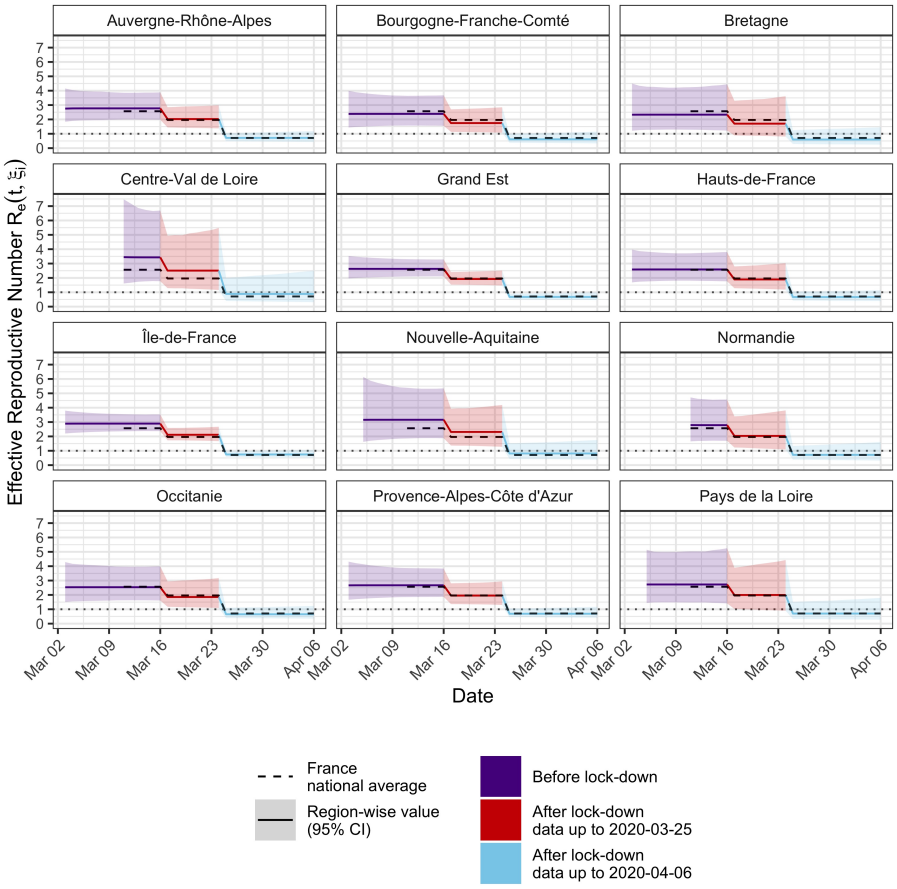

Figure 4: Region specific $R_e(t, \xi_i)$ compared to the national average. Lock-down started on March 17th.

	Region	Minimal confinement duration τ_i^{opt} in days	Date t_i^{ICU} when ICU capacities reached	% of ICU beds occupied at peak	Number of Deaths $0.005 \times R(\infty, \xi_i)$	Final attack rates $R(\infty, \xi_i)/N_i$
$K_2 = 3$	*Auvergne-Rhône-Alpes*	348 [233; 435]	2020-04-10	107% [88%; 127%]	18,567 [15,022; 23,399]	4.6% [3.7%; 5.8%]
	Bourgogne-Franche-Comté	260 [173; 335]	2020-03-25	200% [155%; 254%]	4,045 [3,135; 5,388]	2.9% [2.3%; 3.9%]
	Bretagne	228 [156; 286]		62% [41%; 88%]	5,080 [3,760; 6,891]	3.0% [2.3%; 4.1%]
	Centre-Val de Loire	626 [425; 736]		93% [45%; 172%]	20,532 [11,263; 33,550]	16.0% [8.8%; 26.2%]
	Grand Est	366 [256; 447]	2020-03-26	250% [201%; 307%]	26,710 [20,942; 34,837]	9.7% [7.6%; 12.6%]
	Hauts-de-France	259 [184; 314]		61% [48%; 76%]	16,211 [13,396; 19,767]	5.4% [4.5%; 6.6%]
	Île-de-France	518 [358; 612]	2020-03-30	205% [164%; 261%]	70,331 [52,721; 98,272]	11.5% [8.6%; 16.0%]
	Nouvelle-Aquitaine	848 [521; 990]	2020-04-13	155% [82%; 273%]	32,511 [16,105; 58,929]	10.8% [5.4%; 19.6%]
	Normandie	293 [184; 383]		41% [27%; 57%]	7,643 [5,546; 10,639]	4.6% [3.4%; 6.4%]
	Occitanie	357 [219; 475]		62% [44%; 85%]	9,850 [6,861; 14,850]	3.3% [2.3%; 5.0%]
	Provence-Alpes-Côte d'Azur	411 [240; 546]	2020-04-05	129% [92%; 172%]	10,524 [7,084; 15,864]	4.2% [2.8%; 6.3%]
	Pays de la Loire	373 [232; 478]	2020-04-12	106% [66%; 158%]	13,078 [8,404; 21,423]	6.9% [4.4%; 11.3%]
$K_2 = 5$	*Auvergne-Rhône-Alpes*	148 [119; 170]		79% [67%; 93%]	9,185 [8,096; 10,529]	2.3% [2.0%; 2.6%]
	Bourgogne-Franche-Comté	124 [98; 145]	2020-03-25	158% [124%; 197%]	2,269 [1,857; 2,727]	1.6% [1.3%; 2.0%]
	Bretagne	122 [97; 142]		52% [35%; 73%]	3,229 [2,553; 4,063]	1.9% [1.5%; 2.4%]
	Centre-Val de Loire	167 [118; 206]		38% [24%; 55%]	3,222 [2,405; 4,289]	2.5% [1.9%; 3.4%]
	Grand Est	156 [126; 181]	2020-03-26	178% [148%; 210%]	12,307 [10,584; 14,286]	4.5% [3.8%; 5.2%]
	Hauts-de-France	136 [112; 156]		51% [40%; 62%]	10,072 [8,755; 11,538]	3.4% [2.9%; 3.9%]
	Île-de-France	181 [144; 209]	2020-03-30	125% [105%; 147%]	23,331 [20,019; 27,317]	3.8% [3.3%; 4.4%]
	Nouvelle-Aquitaine	175 [126; 211]		65% [44%; 87%]	4,337 [3,172; 5,657]	1.4% [1.1%; 1.9%]
	Normandie	127 [98; 151]		31% [21%; 42%]	3,933 [3,211; 4,818]	2.4% [1.9%; 2.9%]
	Occitanie	141 [110; 166]		45% [33%; 61%]	4,553 [3,597; 5,782]	1.5% [1.2%; 2.0%]
	Provence-Alpes-Côte d'Azur	146 [113; 171]		88% [68%; 112%]	4,232 [3,423; 5,231]	1.7% [1.4%; 2.1%]
	Pays de la Loire	146 [111; 172]		74% [49%; 105%]	5,717 [4,166; 7,732]	3.0% [2.2%; 4.1%]
$K_2 = 10$	*Auvergne-Rhône-Alpes*	102 [87; 114]		72% [60%; 84%]	7,219 [6,431; 8,034]	1.8% [1.6%; 2.0%]
	Bourgogne-Franche-Comté	89 [75; 101]	2020-03-25	144% [115%; 176%]	1,838 [1,547; 2,167]	1.3% [1.1%; 1.6%]
	Bretagne	90 [75; 102]		48% [32%; 67%]	2,701 [2,147; 3,292]	1.6% [1.3%; 2.0%]
	Centre-Val de Loire	93 [74; 109]		30% [19%; 45%]	2,079 [1,631; 2,615]	1.6% [1.3%; 2.0%]
	Grand Est	105 [90; 118]	2020-03-26	157% [131%; 186%]	9,426 [8,149; 10,859]	3.4% [3.0%; 3.9%]
	Hauts-de-France	99 [85; 111]		47% [38%; 57%]	8,353 [7,392; 9,444]	2.8% [2.5%; 3.2%]
	Île-de-France	114 [98; 128]	2020-03-31	107% [92%; 124%]	16,766 [14,859; 19,112]	2.7% [2.4%; 3.1%]
	Nouvelle-Aquitaine	101 [82; 116]		53% [37%; 72%]	2,856 [2,168; 3,665]	1.0% [0.7%; 1.2%]
	Normandie	87 [71; 100]		28% [19%; 38%]	3,104 [2,588; 3,721]	1.9% [1.6%; 2.3%]
	Occitanie	97 [80; 110]		41% [30%; 54%]	3,538 [2,825; 4,463]	1.2% [1.0%; 1.5%]
	Provence-Alpes-Côte d'Azur	97 [81; 111]		78% [59%; 98%]	3,202 [2,576; 3,901]	1.3% [1.0%; 1.5%]
	Pays de la Loire	98 [79; 111]		66% [43%; 92%]	4,361 [3,158; 5,763]	2.3% [1.7%; 3.0%]

Table 6: Predicted outcomes under lockdown during optimal time τ_i^{opt}. Note that ICU and Deaths are not directly modeled as a compartment in our model and based on crude estimates. ICU capacities are considered as of 2018-12-31 according to DREES (2019) and do not account for the recent surge in ICU beds in France.

Region	**Attack rates: Infected proportion of the population in % computed as (E+I+A+H+R)/N**				
	2020-03-17	**2020-04-13**	**2020-05-11**	**2020-06-08**	**2020-06-22**
Auvergne-Rhône-Alpes	0.6 [0.5; 0.6]	2.5 [2.2; 2.8]	3.0 [2.6; 3.4]	3.1 [2.6; 3.6]	3.1 [2.6; 3.6]
Bourgogne-Franche-Comté	0.5 [0.4; 0.6]	1.8 [1.5; 2.1]	2.0 [1.7; 2.5]	2.1 [1.7; 2.6]	2.1 [1.7; 2.6]
Bretagne	0.6 [0.5; 0.8]	2.1 [1.6; 2.6]	2.3 [1.8; 3.0]	2.4 [1.8; 3.1]	2.4 [1.8; 3.1]
Centre-Val de Loire	0.3 [0.3; 0.4]	2.7 [2.0; 3.6]	3.9 [2.8; 5.5]	4.6 [3.2; 6.6]	4.8 [3.3; 6.9]
Grand Est	1.0 [0.9; 1.2]	4.9 [4.2; 5.7]	5.9 [5.0; 7.0]	6.2 [5.2; 7.4]	6.2 [5.2; 7.4]
Hauts-de-France	1.1 [1.0; 1.2]	3.7 [3.2; 4.2]	4.1 [3.5; 4.8]	4.2 [3.6; 4.9]	4.2 [3.6; 4.9]
Île-de-France	0.7 [0.6; 0.8]	4.1 [3.6; 4.8]	5.4 [4.6; 6.4]	5.8 [4.9; 7.0]	5.9 [5.0; 7.1]
Nouvelle-Aquitaine	0.2 [0.2; 0.3]	1.6 [1.2; 2.0]	2.2 [1.6; 3.1]	2.6 [1.8; 3.7]	2.7 [1.8; 3.8]
Normandie	0.6 [0.5; 0.7]	2.6 [2.1; 3.2]	3.1 [2.4; 3.9]	3.2 [2.5; 4.0]	3.2 [2.5; 4.1]
Occitanie	0.4 [0.3; 0.5]	1.7 [1.3; 2.1]	2.0 [1.5; 2.6]	2.1 [1.6; 2.7]	2.1 [1.6; 2.7]
Provence-Alpes-Côte d'Azur	0.4 [0.3; 0.4]	1.8 [1.5; 2.3]	2.3 [1.8; 2.9]	2.4 [1.9; 3.1]	2.4 [1.9; 3.1]
Pays de la Loire	0.7 [0.5; 0.9]	3.3 [2.3; 4.5]	4.0 [2.8; 5.7]	4.2 [2.9; 6.0]	4.2 [2.9; 6.0]
France	0.6 [0.5; 0.7]	2.9 [2.4; 3.5]	3.6 [2.9; 4.4]	3.8 [3.1; 4.7]	3.8 [3.1; 4.8]

Table 7: Model predictions for the proportion of Infected and Immunized in the population (deaths not taken into account), assuming continued lockdown until then.

lockdown lift on May 11th 2020 We simulated the effect of lifting the lockdown on May 11th 2020 assuming that after this date the transmission goes back to its value before lockdown. Figure 5 shows the predicted dynamics for each region. In every region, we observed a large rebound occurring either in June or July. The timing and magnitude of this rebound is largely influenced by the importance of the first wave, that is successfully contained thanks to the lockdown. These results strongly argue for enforcing other NPIs when lockdown is lifted in order to contain R_e below 1 and prevent this predictable rebound of the epidemic.

4 Discussion

In this work, we provide estimations of the key parameters of the dynamics of the COVID-19 epidemic in French regions as well as forecasts according to NPIs especially regarding the proportion of infected when lifting the lockdown policy.

The point estimates of the basic reproductive ratios for French regions fluctuated between 2.4 and 3.4 before lockdown took effect, but according to the uncertainty around these estimates they are not substantially different from one region to another. Therefore, observed differences in the number of cases were due to the epicemic starting first in *Grand Est* and *Île-de-France* regions. These estimates were close to those reported before isolation using other models (Alizon et al., 2020; Flaxman et al., 2020). The model provided estimates of the impact of the lockdown on the effective reproductive ratio and although recent data led to a substantial reduction of R_e after the lockdown, it remains close to 1 thus without a clear extinction of the epidemic. These estimates should be updated with more recent data that may lead to an estimated R below 1. On the other hand, it is an argument to add other measures such as intensive testing and strict isolation of cases. In addition, the model provides estimates of the size of the population of people who have been or are currently infected. As already reported (Di Domenico et al., 2020a), this proportion of subjects is around 2 to 4 percent, so excluding any herd immunity and control of the epidemic by having a large proportion of people already infected and therefore not susceptible. With our estimates of basic reproduction ratio, the epidemic would become extinct by herd immunity with a proportion of 89.5% (95% CI [88.0%; 90.7%]) of infected people.

Interpretation of our results is conditional on the mechanistic model illustrated in Figure 1, and careful attention must be given to the parameters set

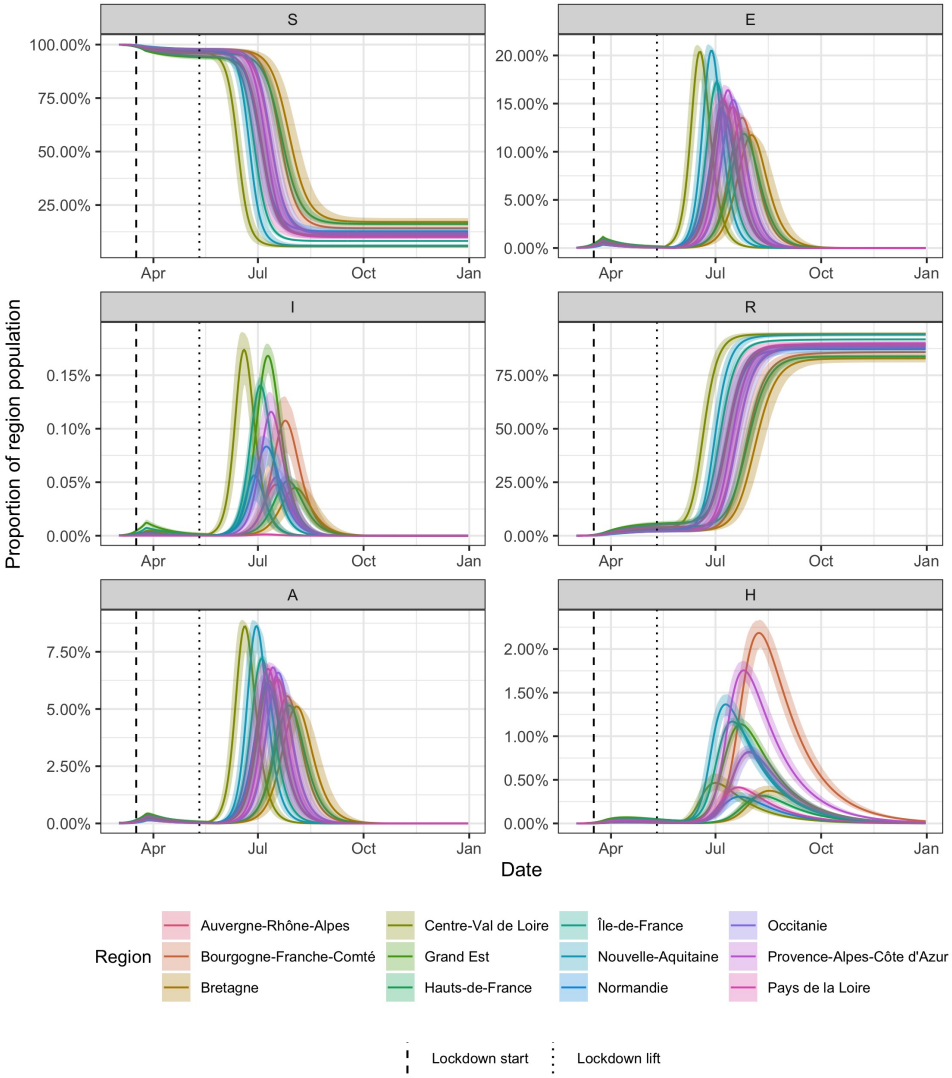

Figure 5: Region specific predicted dynamics if lockdown is entirely lifted on May 11th 2020.

from the scientific literature and detailed in Table 2, as updated estimates are published every day. First and foremost, our model takes only two kinds of infectious cases into account: confirmed cases I, and unascertained cases A. Our observation model takes I as the number of infectious cases confirmed by a positive PCR SARS-Cov-2 test. Thus, A can be interpreted as

unconfirmed symptomatic cases that can be diagnosed by a GP visit (possibly through remote teleconsultation). This is a very simple representation of the COVID-19 infection, which can have various degrees of severity (e.g. asymptomatic, mild, severe) that could be themselves modeled into different compartments. However, very little data is currently available to gather sufficient information to be able to distinguish between those infectious states. Second, our model does not have a compartment for COVID-19 patients in ICU, and the number of occupied ICU beds is simply taken as a fixed percentage (25% based on an estimate from the Bordeaux *CHU* University Hospital). Meanwhile ICU bed capacity does not account for the recent surge of available ICU beds in response to the COVID-19 epidemic. Compared to Wang et al. (2020), our model does not feature an inflow of susceptibles n (and matching outflow) but population movement across regions are limited during the isolation period (see Appendix A3 for a thorough discussion). Deaths were also not distinguished from recoveries in the R compartment, but over the observation period this did not impact the main estimates. Third, our model does not take into account the age-structure of the population on the contrary to the recently posted report Salje et al. (2020) using French data. Interestingly, although the models were different and the data not fully identical, our results were comparable. Actually, our approach captures a part of the unexplained variability between regions through the random effects. This variability might be explained at least partly through the difference in age-structure and probability of hospitalization according to the age.

We would like to underline the interest of making the data publicly accessible as done by *Santé Publique France* on the data.gouv.fr web portal, hence allowing our group to work immediately on the topics. Furthermore, we have made our code fully available on GitHub `www.github.com/sistm/SEIRcovid19` and we are currently working on packaging our software for facilitating its dissemination and re-use.

In conclusion, the lockdown has clearly helped controlling the epidemics in France in every region. The number of infected people varies from one region to the other because of the variations in the epidemic start in these regions (both in terms of timing and size). Hence, the predicted proportion of infected people as of May 11 varies, but stays below 10 % everywhere. It is clear from this model, as in other published models (Di Domenico et al., 2020b; Flaxman et al., 2020), that a full and instantaneous lockdown lift would lead to a rebound. Additional measures may help in controlling the number of new infections such as strict case isolation, contact tracing (Di Domenico et al., 2020b) and certainly a protective vaccine for which the

strategy of administration to the population remains to be defined (Amanat and Krammer, 2020; Lurie et al., 2020; Thanh et al., 2020).

Availability

The data from the SurSaUD® database regarding COVID-19 is available from the `data.gouv` French government platform at https://www.data.gouv.fr/fr/datasets/donnees-des-urgences-hospitalieres-et-de-sos-medecins-relatives-a-lepidemie-de-covid-19. The source code used for this work is available on GitHub at www.github.com/sistm/SEIRcovid19.

Acknowledgements

The authors thank Romain Greffier for his time in discussing the aggregated features of COVID-19 patients care at the Bordeaux University Hospital. The authors thank the `opencovid-19` initiative for their contribution in opening the data used in this article. BPH thanks Vincent Pey for discussions about the clinical characteristics of the COVID-19 infection. This work is supported in part by the GESTEPID Inria Mission COVID19.

Competing interest

The authors have no competing interests to declare.

Author contributions

MP, LW, RT and BPH designed the study. MP and BPH analyzed the data. MP, DD and BPH implemented the software code. QC performed identifiability and asymptotic analysis of the model. MP, LW, RT and BPH interpreted the results and wrote the manuscript.

References

Alizon, S., Bénéteau, T., Choisy, M., Danesh, G., Djidjou-Demasse, R., Elie, B., Michalakis, Y., Reyné, B., Richard, Q., Selinger, C., and Sofonea, M. T. (2020). Rapport #1 : Estimation du nombre de reproduction de l'épidémie de covid-19 en france.

Amanat, F. and Krammer, F. (2020). Sars-cov-2 vaccines: status report. *Immunity*.

Bellu, G., Saccomani, M. P., Audoly, S., and D'Angiò, L. (2007). Daisy: A new software tool to test global identifiability of biological and physiological systems. *Computer methods and programs in biomedicine*, 88(1):52–61.

Bernard-Stoecklin, S., Rolland, P., Silue, Y., Mailles, A., Campese, C., Simondon, A., Mechain, M., Meurice, L., Nguyen, M., Bassi, C., Yamani, E., Behillil, S., Ismael, S., Nguyen, D., Malvy, D., Lescure, F. X., Georges, S., Lazarus, C., Tabaï, A., Stempfelet, M., Enouf, V., Coignard, B., Levy-Bruhl, D., and team, I. (2020). First cases of coronavirus disease 2019 (covid-19) in france: surveillance, investigations and control measures, january 2020. *Eurosurveillance*, 25(6).

Caserio-Schönemann, C., Bousquet, V., Fouillet, A., and Henry, V. (2014). The French syndromic surveillance system SurSaUD®. *Bulletin Epidémiologique Hebdomadaire*, 3–4:38–44.

Delyon, B., Lavielle, M., and Moulines, E. (1999). Convergence of a Stochastic Approximation Version of the EM Algorithm. *The Annals of Statistics*, 27(1):94–128.

Di Domenico, L., Pullano, G., Coletti, P., Hens, N., and Colizza, V. (2020a). Report #8: Expected impact of school closure and telework to mitigate covid-19 epidemic in france.

Di Domenico, L., Pullano, G., Sabbatini, C. E., Boëlle, P.-Y., and Colizza, V. (2020b). Report #9: Expected impact of lockdown in île-de-france and possible exit strategies.

DREES (2019). Statistique annuelle des établissements de santé. accessed: 2020-03-25.

Fang, Y., Nie, Y., and Penny, M. (2020). Transmission dynamics of the covid-19 outbreak and effectiveness of government interventions: A data-driven analysis. *Journal of medical virology*.

Ferguson, N., Laydon, D., Nedjati Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunuba Perez, Z., Cuomo-Dannenburg, G., , Dighe, A., Dorigatti, I., Fu, H., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Okell, L. C., van Elsland, S., Thompson, H., Verity, R., Volz, E., Wang, H., Wang, Y., Walker, P. G., Walters, C., Winskill, P.,

Whittaker, C., Donnelly, C. l. A., Riley, S., and Ghani, A. C. (2020). Report 9: Impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand.

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Coupland, H., Mellan, T. A., Zhu, H., Berah, T., Eaton, J. W., Guzman, P. N. P., Schmit, N., Cilloni, L., Ainslie, K. E. C., Baguelin, M., Blake, I., Boonyasiri, A., Boyd, O., Cattarino, L., Ciavarella, C., Cooper, L., Cucunubá, Z., Cuomo-Dannenburg, G., Dighe, A., Djaafara, B., Dorigatti, I., van Elsland, S., FitzJohn, R., Fu, H., Gaythorpe, K., Geidelberg, L., Grassly, N., Green, W., Hallett, T., Hamlet, A., Hinsley, W., Jeffrey, B., Jorgensen, D., Knock, E., Laydon, D., Nedjati-Gilani, G., Nouvellet, P., Parag, K., Siveroni, I., Thompson, H., Verity, R., Volz, E., Walters, C., Wang, H., Wang, Y., Watson, O., Winskill, P., Xi, X., Whittaker, C., Walker, P. G. T., Ghani, A., Donnelly, C. A., Riley, S., Okell, L. C., Vollmer, M. A. C., Ferguson, N. M., and Bhatt, S. (2020). Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on covid-19 in 11 european countries. *Imperial College COVID-19 Response Team*.

Fox, G. J., Trauer, J. M., and McBryde, E. (2020). Modelling the impact of covid-19 upon intensive care services in new south wales. *The Medical Journal of Australia*, 212(10):1.

Gane, S. B., Kelly, C., and Hopkins, C. (2020). Isolated sudden onset anosmia in covid-19 infection. a novel syndrome? *Rhinology*.

Grasselli, G., Pesenti, A., and Cecconi, M. (2020). Critical Care Utilization for the COVID-19 Outbreak in Lombardy, Italy: Early Experience and Forecast During an Emergency Response. *JAMA*.

Greenhalgh, T., Koh, G. C. H., and Car, J. (2020). Covid-19: a remote assessment in primary care. *BMJ*, 368.

Guan, W.-j., Ni, Z.-y., Hu, Y., Liang, W.-h., Ou, C.-q., He, J.-x., Liu, L., Shan, H., Lei, C.-l., Hui, D. S., Du, B., Li, L.-j., Zeng, G., Yuen, K.-Y., Chen, R.-c., Tang, C.-l., Wang, T., Chen, P.-y., Xiang, J., Li, S.-y., Wang, J.-l., Liang, Z.-j., Peng, Y.-x., Wei, L., Liu, Y., Hu, Y.-h., Peng, P., Wang, J.-m., Liu, J.-y., Chen, Z., Li, G., Zheng, Z.-j., Qiu, S.-q., Luo, J., Ye, C.-j., Zhu, S.-y., and Zhong, N.-s. (2020). Clinical characteristics of coronavirus disease 2019 in china. *New England Journal of Medicine*.

INED (2020). Insee: Recensements de population, estimations de population.

Kraemer, M. U. G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D. M., , du Plessis, L., Faria, N. R., Li, R., Hanage, W. P., Brownstein, J. S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O. G., and Scarpino, S. V. (2020). The effect of human mobility and control measures on the covid-19 epidemic in china. *Science*.

Kuhn, E. and Lavielle, M. (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics & Data Analysis*, 49(4):1020–1038.

Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., Azman, A. S., Reich, N. G., and Lessler, J. (2020). The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine*.

Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K. S. M., Lau, E. H. Y., Wong, J. Y., Xing, X., Xiang, N., Wu, Y., Li, C., Chen, Q., Li, D., Liu, T., Zhao, J., Liu, M., Tu, W., Chen, C., Jin, L., Yang, R., Wang, Q., Zhou, S., Wang, R., Liu, H., Luo, Y., Liu, Y., Shao, G., Li, H., Tao, Z., Yang, Y., Deng, Z., Liu, B., Ma, Z., Zhang, Y., Shi, G., Lam, T. T. Y., Wu, J. T., Gao, G. F., Cowling, B. J., Yang, B., Leung, G. M., and Feng, Z. (2020a). Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *New England Journal of Medicine*, 382(13):1199–1207.

Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., and Shaman, J. (2020b). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (sars-cov2). *Science*.

Lixoft SAS (2019). Monolix v2019r2.

Lurie, N., Saville, M., Hatchett, R., and Halton, J. (2020). Developing covid-19 vaccines at pandemic speed. *New England Journal of Medicine*.

Magal, P. and Webb, G. (2020). Predicting the number of reported and unreported cases for the covid-19 epidemic in south korea, italy, france and germany. *medRxiv*.

Massonnaud, C., Roux, J., and Crépey, P. (2020). Covid-19: Forecasting short term hospital needs in france. *medRxiv*.

Prem, K., Liu, Y., Russell, T. W., Kucharski, A. J., Eggo, R. M., Davies, N., for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, C., Jit, M., and Klepac, P. (2020). The effect of control

strategies to reduce social mixing on outcomes of the covid-19 epidemic in wuhan, china: a modelling study. *Lancet Public Health*.

Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., and Timmer, J. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15):1923–1929.

Roques, L., Klein, E., Papaix, J., Sar, A., and Soubeyrand, S. (2020). Using early data to estimate the actual infection fatality ratio from covid-19 in france. *medRxiv*.

Salje, H., Tran Kiem, C., Lefrancq, N., Courtejoie, N., Bosetti, P., Paireau, J., Andronico, A., Hoze, N., Richet, J., Dubost, C.-L., Le Strat, Y., Lessler, J., Bruhl, D. L., Fontanet, A., Opatowski, L., Boëlle, P.-Y., and Cauchemez, S. (2020). Estimating the burden of SARS-CoV-2 in France. working paper or preprint.

Soetaert, K., Petzoldt, T., Setzer, R. W., and Petzoldt, M. T. (2020). Package 'desolve'. *Writing Code in Compiled Languages*, pages 1–51.

Tang, B., Xia, F., Tang, S., Bragazzi, N. L., Li, Q., Sun, X., Liang, J., Xiao, Y., and Wu, J. (2020). The effectiveness of quarantine and isolation determine the trend of the covid-19 epidemics in the final phase of the current outbreak in china. *International Journal of Infectious Diseases*.

Thanh, L. T., Andreadakis, Z., Kumar, A., Gómez, R. R., Tollefsen, S., Saville, M., and Mayhew, S. (2020). The covid-19 vaccine development landscape. *Nature reviews. Drug discovery*.

Tian, H., Liu, Y., Li, Y., Wu, C.-H., Chen, B., Kraemer, M. U. G., Li, B., Cai, J., Xu, B., Yang, Q., Wang, B., Yang, P., Cui, Y., Song, Y., Zheng, P., Wang, Q., Bjornstad, O. N., Yang, R., Grenfell, B. T., Pybus, O. G., and Dye, C. (2020). An investigation of transmission control measures during the first 50 days of the covid-19 epidemic in china. *Science*.

Valleron, A.-J., Bouvet, E., Garnerin, P., Ménares, J., Heard, I., Letrait, S., and Lefaucheux, J. (1986). A computer network for the surveillance of communicable diseases: the french experiment. *American journal of public health*, 76(11):1289–1292.

Wang, C., Liu, L., Hao, X., Guo, H., Wang, Q., Huang, J., He, N., Yu, H., Lin, X., Pan, A., Wei, S., and Wu, T. (2020). Evolving epidemiology and

impact of non-pharmaceutical interventions on the outbreak of coronavirus disease 2019 in wuhan, china. *medRxiv*.

Wilder, B., Charpignon, M., Killian, J. A., Ou, H.-C., Mate, A., Jabbari, S., Perrault, A., Desai, A., Tambe, M., and Majumder, M. S. (2020). The role of age distribution and family structure on covid-19 dynamics: A preliminary modeling assessment for hubei and lombardy (march 31, 2020). *Available at SSRN: 3564800*.

World Health Organization (2020a). Coronavirus disease 2019 (covid-19) situation report – 84. accessed: 2020-04-14.

World Health Organization (2020b). Naming the coronavirus disease (covid-19) and the virus that causes it. accessed: 2020-04-02.

World Health Organization (2020c). Who director-general's opening remarks at the media briefing on covid-19 – 11 march 2020. accessed: 2020-04-02.

Wu, Z. and McGoogan, J. M. (2020). Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72,314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA*, 323(13):1239–1242.

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G. F., and Tan, W. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine*, 382(8):727–733.

Appendix

A1 Data sources and reporting variability

One of the difficulties in modeling the COVID-19 epidemic in real-time is the reliability and comparability of data sources an reporting artifacts.

Differences bewteen the SurSaUD® database and *Santé publique France* reports The *Santé Publique France* epidemiological reports cite *SI-VIC* as their source. *SI-VIC* is a special system that can be activated for the identification and monitoring of victims from terror attacks and exceptional health situations, and that was activated on March 13th 2020 for the COVID-19 epidemic and is maintained by the French *Agence du Numérique en Santé*. Because SI-VIC is a declarative web-based plateform (with data populated by regional health agencies, emergency services and hospitals) it was likely missing some case declaration when first launched. On the contrary, the SurSaUD® database. The Figure S1 illustrate the discrepencies and differences between the two database.

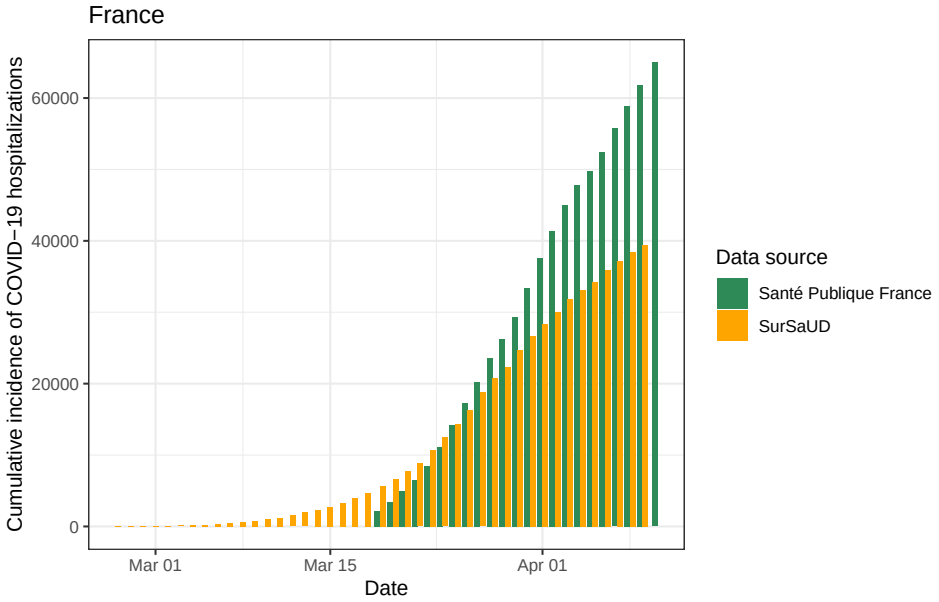

Figure S1: Cumulative incidence hospitalization for COVID-19 at France national level according to either *Santé publique France* or the SurSaUD® database

Of note, the earlier slowing down of emergency admissions for COVID-19 suspicions observed in the SurSaUD® data compared to the *SI-VIC* data could lead to an optimistic biais regarding the epidemic stage and evolution in our estimations and predictions . However, since the *SI-VIC* data are not publicly accessible before March 18th, we are currently unable to use this data source to estimate the impact of the lockdown.

Under-reporting or over-estimation of COVID-19 deaths According to Roques et al. (2020), the Infection Fatality Ratio (IFR) for COVID-19 is 0.5% (95%-CI: 0.3; 0.8). Figure S2 shows that using this IFR over-estimate compared to the death currently reported by *Santé Publique France*.

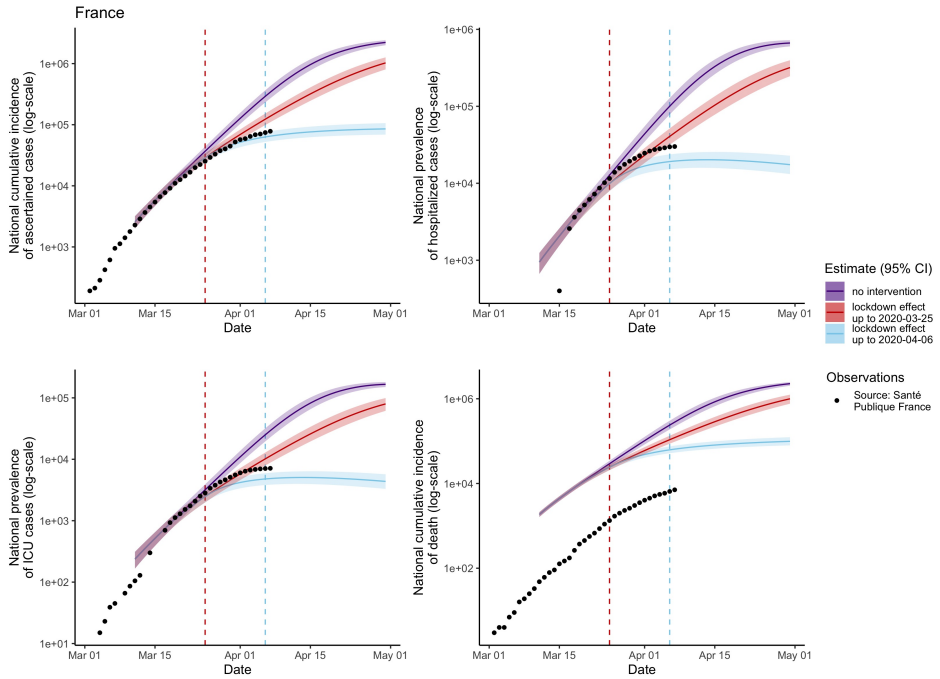

Figure S2: Observed incident number of ascertained cases and hospitalization at France national level compared to predicted ones, based on estimations with data collected up to either March 16 (before lockdown), March 25th or April 6th (delimited by the vertical line).

A2 Theoretical identifiability analysis of the $SEIRAH$ model

Based on ODE model (1), the statistical analysis of the population evolution can be turned into a parameter estimation problem. Theoretical identifiability (i.e. the possibility of learning the true values of the model parameters) of epidemiological compartment models is rarely checked. Although it relies on unrealistic hypothesis (namely that incident number of ascertained cases and hospitalized cases are observed continuously, without noise and for an infinite length of time in our case), this framework provides important guarantees for the results. To determine which parameters from Table 2 can be accurately identified from the available observations we first evaluate the identifiability of this SEIRAH structure with the DAISY software from Bellu et al. (2007) (based on differential algebra results). We conclude that there is global identifiability of $\widetilde{\xi} = (b, D_q, r, D_e, D_I)$ with known α and D_h, even if initial conditions are unknown – which is our case for (E_0, A_0). Practical identifiability (for which most existing evaluation methods rely on estimations and are based on Fisher information matrix or likelihood profiling (Raue et al., 2009)) of these parameters will be discussed in section 2.3.

Figure S3 present the occupancy numbers of the six compartments of the model when one parameter is varied while the others are fixed. Because the epidemics start date and state is different in each region, it is necessary to estimate in priority E_0 and A_0. b, D_e and D_I have a similar impact on the simulations, and we chose to prioritize the estimation of the transmission b as it is a important actionable driver of the epidemics, that can potentially be reduced by interventions. r and D_q have little impact on $I^{(in)}$ but are informative for $H^{(in)}$. Since r can be estimated from other data sources, we prioritized the estimation of D_q. In addition, there is a strong rational to estimate b and D_q as they are directly involved in the computation of the reproductive number. So in the end, we assume D_e, D_I, and r fixed and we will estimate E_0, A_0, b, and D_q. Both D_e and D_I are set to a common value across regions, estimated from previous studies referenced in Table 2. To set r, we used data collected from general practitioners through the re-purposing of the *Réseau Sentinelles* network to monitor COVID-19 and provide a weekly estimation of the number of incident symptomatic cases (regardless of their confirmation status through a PCR test) at the region level. Thus, for each region, r_i is set to r_s the ratio of observed incident number of ascertained cases over the incident number of symptomatic cases (as defined by the incident symptomatic cases reported by the *Réseau Sentinelles* network). Values are provided in Table 1.

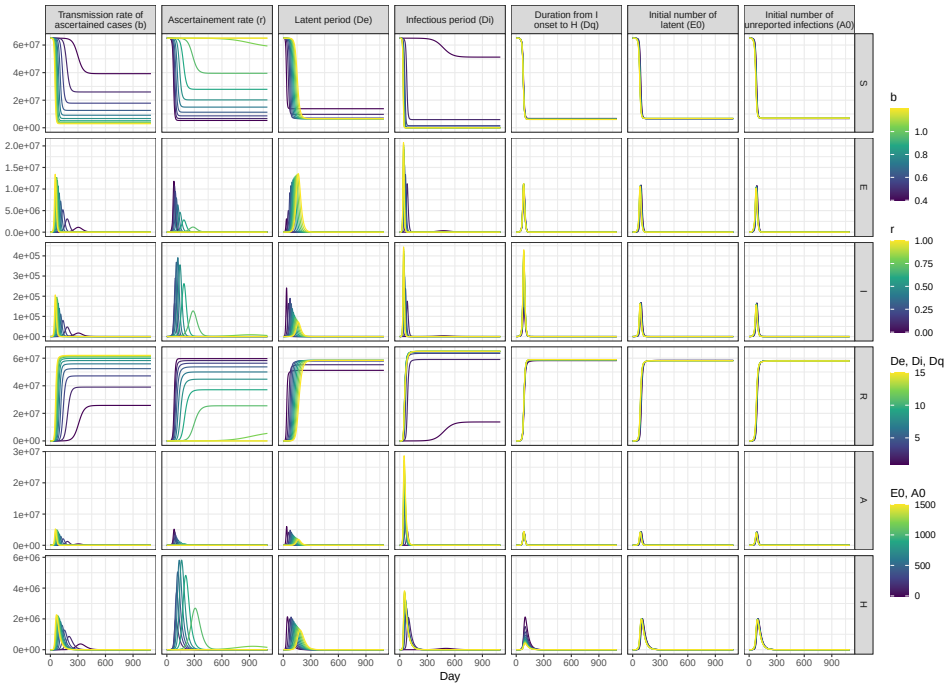

Figure S3: Impact of parameters on the compartment dyamics. Parameters fixed values (when other parameters are varied): $b = 0.8$, $r = 0.1$, $D_e = 5.1$, $D_q = 1.2$, $E_0 = 730$, and $A_0 = 100$

A3 SEIRAH asymptotics with sustained susceptible inflow n

The model ((1)) is a simplified version of the ODE system presented in Wang et al. (2020):

$$\begin{cases} \dfrac{dS}{dt} &= -\dfrac{bS(I+\alpha A)}{N} - \dfrac{nS}{N-I-H} + n \\ \dfrac{dE}{dt} &= \dfrac{bS(I+\alpha A)}{N} - \dfrac{E}{D_e} - \dfrac{nE}{N-I-H} \\ \dfrac{dI}{dt} &= \dfrac{rE}{D_e} - \dfrac{I}{D_q} - \dfrac{I}{D_I} \\ \dfrac{dR}{dt} &= \dfrac{(I+A)}{D_I} + \dfrac{H}{D_h} - \dfrac{nR}{N-I-H} \\ \dfrac{dA}{dt} &= \dfrac{(1-r)E}{D_e} - \dfrac{A}{D_I} - \dfrac{nA}{N-I-H} \\ \dfrac{dH}{dt} &= \dfrac{I}{D_q} - \dfrac{H}{D_h} \end{cases} \tag{9}$$

In the ODE system (9), Wang et al. (2020) consider the possibility of an exogenous flow of susceptible modeled by a sustained susceptible inflow n. While this make sense for modeling an outbreak in a region surrounded by other regions free of the epidemic (such as the first outbreak of COVID-19 in Wuhan (Hubei, China) in late 2019 - early 2020), it is not suitable for a pandemic where the pathogen is circulating in all regions, which is the current situation in all of France COVID-19.

Nonetheless we study the possible asymptotic steady states of this model, i.e. the constant values $\left(S^{(0)}, E^{(0)}, I^{(0)}, R^{(0)}, A^{(0)}, H^{(0)}\right)$ possibly reached by the system when $t \to \infty$. Without loss of generality, we assume the initial number of removed subjects is set to $R(t=0) = 0$. By definition, the steady

states verify:

$$\begin{cases} 0 &= -\dfrac{bS^{(0)}(I^{(0)}+\alpha A^{(0)})}{N} - \dfrac{nS^{(0)}}{N-(1+\beta_2)I^{(0)}} + n \\ 0 &= \dfrac{bS(I^{(0)}+\alpha A^{(0)})}{N} - \dfrac{\beta_1 I^{(0)}}{D_e} - \dfrac{n\beta_1 I^{(0)}}{N-(1+\beta_2)I^{(0)}} \\ E^{(0)} &= \beta_1 I^{(0)} \\ 0 &= \dfrac{(I^{(0)}+A^{(0)})}{D_I} + \dfrac{\beta_2 I^{(0)}}{D_h} - \dfrac{nR^{(0)}}{N-(1+\beta_2)I^{(0)}} \\ 0 &= \dfrac{(1-r)\beta_1 I^{(0)}}{D_e} - \dfrac{A^{(0)}}{D_I} - \dfrac{nA^{(0)}}{N-(1+\beta_2)I^{(0)}} \\ H^{(0)} &= \beta_2 I^{(0)} \end{cases} \tag{10}$$

with $\beta_1 = D_e/r\,(1/D_q + 1/D_i)$ and $\beta_2 = D_h/D_q$.

In the particular case of $n = 0$, the fifth equation in (10) can be simplified and imposes $A^{(0)} = D_i(1-r)\beta_1 I^{(0)}/D_e$ which in combination with the third equation leads to $I^{(0)} = 0$ and thus $A^{(0)} = H^{(0)} = E^{(0)} = 0$. Assuming that $A(t) = I(t) = 0$ for all times t and applying the population conservation principle we identify $(N, 0, 0, 0, 0, 0)$ as a steady state. Since the ODE imposes that S is strictly decreasing as long as $S(t) > 0$ and $A(t) > 0$ (or $I(t) > 0$) and that S and A (or I) cannot reach 0 in finite time if $A_0 \neq 0$ (or $I_0 \neq 0$), this steady state $(N, 0, 0, 0, 0, 0)$ is unstable. Having $A_0 \neq 0$ (or $I_0 \neq 0$) leaves us with $S^{(0)} = 0$ as the only possible asymptotic steady state value and since conservation principle imposes $S_0 + E_0 + I_0 + R_0 + A_0 + H_0 = N$, we end up with $R^{(0)} = N$ and we identify $\left(S^{(0)}, E^{(0)}, I^{(0)}, R^{(0)}, A^{(0)}, H^{(0)}\right) = (0, 0, 0, N, 0, 0)$, as the only possible asymptotic steady when $A_0 \neq 0$ (or $I_0 \neq 0$).

But in presence of an inflow of susceptible $n \neq 0$, can we expect an asymptotic behavior with extinction of the epidemic. If we set $A^{(0)} = I^{(0)} = 0$ in (10), the first and fourth equations give us:

$$\begin{cases} n &= \dfrac{nS^{(0)}}{N} \\ 0 &= \dfrac{nR^{(0)}}{N} \end{cases} \tag{11}$$

hence the potential steady state corresponding to the extinction imposes $S^{(0)} = N$, $R^{(0)} = 0$ and $(0, 0, 0, N, 0, 0)$ is no longer a possible steady state point. The only steady state $(N, 0, 0, 0, 0, 0)$ corresponds to the case where no one gets infected nor removed i.e when the epidemic does not start at all.

In this new setting, $A^{(0)}$, $S^{(0)}$ and $R^{(0)}$ are now functions of $I^{(0)}$ and given by:

$$\begin{cases} A^{(0)} &= \dfrac{\beta_4 I^{(0)}(N-(1+\beta_2)I^{(0)})}{nD_i+N-(1+\beta_2)I^{(0)}} \\ S^{(0)} &= \left(1-\dfrac{\beta_1 I^{(0)}}{nD_e}\right)\left(N-(1+\beta_2)I^{(0)}\right)-\beta_1 I^{(0)} \\ R^{(0)} &= \left(\dfrac{I^{(0)}+A^{(0)}}{D_i}+\dfrac{\beta_2 I^{(0)}}{D_h}\right)\dfrac{N-(1+\beta_2)I^{(0)}}{n} \end{cases} \tag{12}$$

with $\beta_4 = (1-r)\beta_1 D_i/D_e$. From algebraic manipulation of (10), we derive that a necessary and sufficient condition for $I^{(0)}$ to constitute a steady set is to be a positive real solution of the fourth-order polynomial equation:

$$\begin{array}{l} bD_e S^{(0)}(nD_i+(\alpha\beta_4+1)\left(N-(1+\beta_2)I^{(0)}\right))(N-(1+\beta_2)I^{(0)}) \\ -\beta_1 N(nD_e+N-(1+\beta_2)I^{(0)})\left(nD_i+N-(1+\beta_2)I^{(0)}\right)=0 \end{array} \tag{13}$$

which does not admit $I^{(0)}=0$ as a solution if

$$\alpha \neq \frac{(\beta_1(nD_e+N)-bD_eN)\,(nD_i+N)}{bN^2(1-r)\beta_1 D_i}.$$

Thus, there are possibly up to four new non trivial steady states when $n \neq 0$.