## An Extended Bayesian Hierarchical Model Predicted SARS-CoV-2 Cases and Reproductions After Vaccination in England

## Jiajing Zha, Xiangdong Liu<sup>\*</sup>

Department of Economics, Jinan University, Guangzhou, Guangdong, China Corresponding author.

#### Abstract

In the UK, one of the worst affected countries by the SARS-CoV-2 global pandemic, the UK government has released a comprehensive vaccination plan for SARS-CoV-2 and will set up more vaccination sites in the coming months to expand the service to more people. We built an extended Bayesian hierarchical prediction model to predict the number of cases and breeding situation after vaccination in the nine districts of England. Based on the population of each region, the number of deaths and the IFR (the infection mortality ratio) for each region were predicted. We found that EAST, NORTHWEST and SOUTHEAST had the largest IFR, and the corresponding death numbers were 29,079, 28,734 and 25,201, respectively. Reproduction number ( $R_t$ ) is expected to drop below 1 in all regions on January 7, January 12, January 16, January 13, January 10, January 17, January 10, January 18 and January 14, respectively. Major vaccine interventions have been effective in reducing transmission in the nine areas of England given the mortality rate of the infected people and epidemiological characteristics of SARS-CoV-2. We also found that vaccination among people aged 70 to 80 had made a significant contribution to reducing transmission of the virus. The model can be extended to forecast the effects of certain interventions in public health emergencies, the effect of preventing the spread of disease, and the effect of different interventions in different age groups to find the best combination of solutions.

Keywords: Bayesian hierarchical model, prediction, vaccination intervention, SARS-CoV-2, Age stratification

## I. Introduction

According to real-time data from Worldometer, the cases' number of confirmed people reached 10,203,700 by 7:00 PM on January 26, 2021, and the total number of deaths reached 2,147,411 by 2,2.14 million[1].

Covid-19 undermines the normal development of global public health. There is an urgent need for vaccines to better control the spread of the epidemic. At present, various departments are rushing to develop the vaccine and expand production after it is approved by regulatory authorities. UK government released its full vaccination plan for the new crown, and will set up more vaccination sites in the coming months to extend the service to more people. More than 2.2 million people in the UK have now received the first dose of the new crown vaccine [2].

T. Lietman et al. [3] used the pre-exposure vaccine model and the post-exposure vaccine model to provide insights into the impact of vaccination on the tuberculosis epidemic. MD. Samsuzzoha et al. [4] uses a sensitivity analysis method based on mathematical and statistical techniques to determine the importance of the parameters of the infectious disease model for vaccination. Kate M. Bubar<sup>s2</sup> used a mathematical model to compare five priority strategies for age stratification. Mohamed El Fatini et al. [5] established a time-delayed stochastic SIR epidemic model with Crowly-Martin function response and vaccination by establishing a random threshold RS as the condition of the stochastic epidemic system. Taking into account the effects of temporary immunization, vaccination, and Levy Jump on disease transmission, Kuangang Fan [6] developed a stochastic model with nonlinear incidence.

## II. Methodology

England is one of countries which been affected severely by the novel Beta-coronavirus SARS-CoV-2. Over 50 thousand of deaths have been reported wordwide, and 90 death cases each 100,000 people on December 12, 2020. The incidence and impact of epidemic diseases are not the same in different regions, but the incidence of elderly people in nursing homes is very high in each region.

On 5th March 2020, the first SARS-CoV-2 death occurred in England [7]. A week later, the government imposed self-quarantine on those with symptoms. [8]. On March 23, the country entered a state of precaution. During the lockdown period, regions had similar levels of transmission (as shown in Figure 1). This is consistent with movement data, which show that movement during the blockade decreased to a level consistent across the country [9].



Figure 1: Trajectory of the UK SARS-CoV-2 epidemic

Due to the rapid spread of the disease and the characteristics of the epidemic, effective interventions are necessary in order to develop policies to prevent the further deterioration of the public emergency. Therefore, we establish an extended Bayesian hierarchical model based on semi-mechanical Bayesian stratified model [10,11]to explore the effects of interventions with the purpose of reducing the transmission rate of SARS-CoV-2 in the UK. In this paper, we forecast the amounts of deaths, infections, and especially the transmission to show how the virus is spreading. Specifically, we estimate the current rate of infection and the impact of different control measures, propose options for further control measures, and predict what might happen in the future.



Figure 2: UK confirmed cases and deaths (as at 30 January 2021) [12] Left: The confirmed cases; Right: the deaths cases.

The infection may be underestimated due to the time lag of detection, so it is difficult to estimate the  $R_t$  of SARS-COV-2[13,14]. The reverse calculation of the quantity of observed deaths and the infections numbers allows an estimate of the course of the epidemic. We introduce a Bayesian mechanism model that takes into account the duration of infection and observed deaths to infer total infections (incidence) and  $R_t$ . Interventions that successfully reduced  $R_t$  to less than 1 were validated. Forecast the number of deaths that would have happened without intervention by simulating a hypothetical counterfactual situation.

By considering the time difference between infections and deaths, and by combining relevant data between countries to calculate reproduction, more timely and accurate predictions can be made. In our model, some epidemiological parameters (such as IFR) are determined and changes in  $R_t$  are assumed to be an immediate response to the intervention. In the current case, the mortality data on which we are based is incomplete, so our conclusions may be systematically biased, which could be further improved in the future.

We estimate that for the nine regions, intervening measures can bring  $R_t$  below 1 (with a greater than 99% probability of  $R_t < 1.0$ ) and bring the out-break under control. We estimate that the results of our study suggest that major drug interventions -- vaccine interventions -- in particular, Vaccination of people between the ages of 70 and 80 has had a major impact on reducing propagated parameters. Continuous intervention should be taken into account to control the infection of SARS-COV-2.

Some epidemiological parameters in the model are fixed, such as the distribution of mortality and the distribution of case fatality rate of infection [15,16]. We performed sensitivity analysis on these parameters. We assume that the change in  $R_t$  is an immediate response to the intervention and that the intervention has similar effects in different countries. At the same time, the effectiveness of these interventions does not diminish over time. Due to the time lag of the test, it is difficult to completely predict the exact value of  $R_t$ . By learning Bayesian analysis during the whole period and considering the confidence interval of  $R_t$ , the predicted result is reasonable [17,18].



Figure 3. The component of the model

2.1 Determination of research scope

| Table 1. The population of the nine regions of England |               |            |  |  |  |
|--|---------------|------------|--|--|--|
| Region Name  | Abbreviations | Population |  |  |  |
| NORTHEAST  | NE            | 2,669,941  |  |  |  |
| NORTHWEST  | NW            | 7,341,196  |  |  |  |
| YORKSHIREANDTHEHUMBER                                  | YH            | 5,502,967  |  |  |  |
| EASTMIDLANDS   | EM            | 4,835,928  |  |  |  |
| WESTMIDLANDS   | WM            | 5,934,037  |  |  |  |
| EAST   | EE            | 6,236,072  |  |  |  |
| LONDON   | LD            | 8,961,989  |  |  |  |
| SOUTHEAST  | SE            | 9,180,135  |  |  |  |
| SOUTHWEST  | SW            | 5,624,696  |  |  |  |
| Total  | Total         | 56,286,961 |  |  |  |

The population of the nine regions of England are shown in the Table 1[19]:

#### 2.2 Death model

The deaths number is more reliable than the confirmed number. Firstly, early deaths caused by SARS-CoV-2 may have been overlooked. Secondly, death reports vary from country to country and from time to time. Thirdly, reporting delays is normal and may be random. To overcome data limitations, we used uniform data sources, such as noise and pooled information, and conducted sensitivity analysis for under-reporting cases to test our conclusions.

Denote  $D_{t,m}$  as daily deaths for days  $t \in 1, ..., n$  and regions  $m \in 1, ..., 9$ . We use a positive real-valued function  $d_{t,m} = E(D_{t,m})$  to model these daily deaths. We assume  $D_{t,m}$  satisfies a negative binomial distribution.

$$D_{t,m} \sim Negative Binomial\left(d_{t,m}, d_{t,m} + \frac{d_{t,m}^{2}}{\Psi}\right)$$
$$\Psi \sim \mathcal{N}(0,5)$$

The expected deaths numbers in a given area on a given day can be calculated from the number of infections that occurred in the previous days. [20].

## 2.3 Infection model

The updating model is related to the susceptibility - infection-recovery model, and the infected individuals' true number C is modeled by a discrete updating process. This method has a strong theoretical basis in the process of random individual counting such as Hawkes and Bellman-Harris processes. [21,22,23,24].

The number of infections to simulate changes over time, we need to specify a has a density of g ( $\tau$ ) when a person is infected to infect other people. The time between we choose the distribution of gamma distribution:

$$g \sim Gamma$$
 (6.5, 0.62).

We assume that all regions are the same [25]. The number of infections  $c_{t,m}$  can be calculated by

$$c_{t,m} = R_{t,m} \sum_{\tau=0}^{t-1} c_{\tau,m} g_{t-\tau}$$

The serial interval can be discretized by

$$g_s = \int_{\tau=s-0.5}^{s+0.5} g(\tau) d\tau \ s = 2,3,\dots$$

Specifically,

$$g_1 = \int_{\tau=0}^{1.5} g(\tau) d\tau$$

2.4 Reproduction number

A function of Google mobility data [11] can model the reproduction number  $R_t$  with time varying. The forecast of 9 regional parameters is used to evaluate whether the interventions currently in place can reduce  $R_t$  below 1.

Denote  $I_{k,t,m}$  as the  $k^{th}$  Google mobility indicator, at the k time for England region m.

$$R_{t,m} = 2R_{0,m} \frac{-\sum_{k=1}^{4} (\alpha_k + \beta_{m,k}) I_{k,t,m}}{\lambda}$$

 $\frac{1}{\lambda}$  satisfy a logistic function,  $\alpha_k$  represent the effects between M regions.  $\beta_{m,k}$  are specific effects in different regions.

Let  $\alpha_k \sim \mathcal{N}(0,0.5)$ ,  $\beta_{m,k} \sim \mathcal{N}(0,\gamma)$ ,  $\gamma \sim \mathcal{N}(0,0.5)$ . The distribution of  $R_{0,m}$  is:

$$R_{0,m} \sim \mathcal{N}(3.28, |\kappa|)$$
  
$$\kappa \sim \mathcal{N}(0, 0.5)$$

All regions have  $\kappa$  and share information about changes in $R_{0,m}$ . Based on YING [26], we set the specific value as 3.28.

2.5 IFR calculation and sensitivity analysis

Differences in the quality of care between regions of the UK are affected by income. As a result, we revised estimates of the severity of infectious diseases (IFR) to take account of income across regions, and we expected to observe intrinsic regional heterogeneity in the state of health outcomes. Although assumptions about the quality of care affect the quantitative prediction of IFR and the incidence of IFR, they are not due to qualitative changes in our herd immunity, and there is currently a lack of strong conclusions[20]. Details are shown in Table 2:

Table 2. The infection mortality ratio (IFR) of the nine regions of England

| S | tate | IFR  | State | IFR  | State | IFR  |
|---|------|------|-------|------|-------|------|
| ] | EE   | 1.1  | SE    | 1.1  | YH    | 1    |
| Ν | W    | 1.1  | WM    | 1    | EM    | 0.78 |
| ] | NE   | 0.73 | LD    | 0.79 | SW    | 0.89 |

#### 2.6 Vaccine intervention

In the current WHO global allocation framework, bioethicists have identified four allocation principles to allocate of scarce resources [27]:

- (A) Maximize the utility of the resource;
- (B) Give priorities to the poorest;
- (C) Equal treatment of individuals;
- (D) Maximization of social benefits.

Considering age stratification associated with SARS-CoV-2 infection, a key to achieving Principles A and B is to target those at highest risk of death from vaccination.

The British government has said it hopes to start offering the new coronary vaccine, made by Pfizer, to high-risk groups from December 1. The vaccine is now awaiting approval from British regulators to ensure it is safe enough to be given to millions of British citizens. Nursing home residents and healthcare workers will be the first to receive the vaccine, followed by frontline staff in hospitals and the rest of the population will be given priority according to age groups. The plan is to have all priority populations vaccinated by the end of February. For the elderly, the specific vaccination plan is as follows:

- 80 + -Starting in mid-December
- everyone between 70 and 80 years old from the end of December
- everyone between the ages of 65 and 70 -- starting in early January 2021
- All high and moderate risk under age 65 starting in early January 2021
- Everyone ages 50 to 65 Starting mid-January 2021

Based on the policy, we choose 4 time points as the nodes of drug intervention to predict the future population changes of England regions. Intervention nodes are shown in Table 3:

| Region | County          | vaccine 80 | vaccine 70 | vaccine 65 | vaccine 55 |
|--------|-----------------|------------|------------|------------|------------|
| EM     | Derbyshire      | 15/12/2020 | 30/12/2020 | 05/01/2021 | 15/01/2021 |
| EE     | Lincoinshire    | 15/12/2020 | 30/12/2020 | 05/01/2021 | 15/01/2021 |
| LD     | Greater London  | 15/12/2020 | 30/12/2020 | 05/01/2021 | 15/01/2021 |
| NE     | Northumberland  | 15/12/2020 | 30/12/2020 | 05/01/2021 | 15/01/2021 |
| NW     | Cumbria         | 15/12/2020 | 30/12/2020 | 05/01/2021 | 15/01/2021 |
| SE     | Oxfordshire     | 15/12/2020 | 30/12/2020 | 05/01/2021 | 15/01/2021 |
| SW     | Gloucestershire | 15/12/2020 | 30/12/2020 | 05/01/2021 | 15/01/2021 |
| WM     | Staffordshire   | 15/12/2020 | 30/12/2020 | 05/01/2021 | 15/01/2021 |
| YH     | Leicestershire  | 15/12/2020 | 30/12/2020 | 05/01/2021 | 15/01/2021 |

Table 3. The vaccine interventions of the nine regions of England in represent county

Among them, vaccine 80 represents the vaccination of the population over 80 years old, 70 represents the vaccination of the population over 70 years old, 65 represents the vaccination of the population over 65 years old, and 55 represents the vaccination of the population over 55 years old. The date comes from the UK government's official release of information. The specific vaccination situation needs to be determined according to the actual situation of each district.

## 2.7 Mobility

Due to the England region, four covariables concerned with Google mobility were included. This is due to the collinearity of the average.

We used mobile report data [28] to evaluate the effect of interventions as time goes by. The result shows the percentage of estimation changes in retail and entertainment venues, parks, grocery stores, pharmacies, etc. We used November 12, 2020 solstice data for December 17.



## **III. Results and discussion**

Figure 4. Estimates of infections, deaths and R<sub>t</sub>. Subfigure 1: Number of infections per day. The dark pink bars represent reported cases, the pink bars represent predicted infections, and the light pink bars represent 95% confidence intervals. Subfigure 2: Daily deaths, where the dark pink bars represent reported deaths and the pink bars represent predicted deaths. Subfigure 3, R<sub>t</sub>, dark blue represents a 50% CI, and light blue represents a 95% CI.

According to our results (shown as in Figure 4) from the extended Bayesian hierarchical prediction model we established, the IFR values of EE, NW and SE reached the maximum. Based on the population of each region, we calculated the death toll of the three regions as: 29079,28734,25201 respectively. The regional reproductive number will be on Jan. 7, Jan. 12, Jan. 16, Jan. 13, Jan. 10, Jan. 17, Jan. 10, Jan. 18, and Jan. 14. Combined with ISSN: 0010-8189

# © CONVERTER 2020 www.converter-magazine.info

awareness of infectious diseases and global prevention measures, national interventions are feasible. Give the observed fatality rate of infection and the epidemiology of SARS-CoV-2, the major vaccine interventions have been effective in reducing transmission in the nine areas of England we considered, vaccination among people aged 70 to 80 has made a significant contribution to reducing transmission of the virus. In our model, we find that these interventions control the prevalence in the current period and can reduce  $R_t$  below 1. If current trends continue, we can contain the epidemic well. All results are shown in Table 4.

| State | IFR  | Pop.      | Deaths | Deaths(ppm) | cases    | case5%  | case95%  |
|-------|------|-----------|--------|-------------|----------|---------|----------|
| EE    | 1.1  | 6,236,072 | 29079  | 4.15E+09    | 31600000 | 1370000 | 61300000 |
| NW    | 1.1  | 7,341,196 | 28734  | 3.59E+09    | 4120000  | 2150000 | 5750000  |
| NE    | 0.73 | 2,669,941 | 25613  | 2.56E+10    | 24700000 | 1200000 | 48300000 |
| SE    | 1.1  | 9,180,135 | 25201  | 2.52E+09    | 2280000  | 1930000 | 2580000  |
| WM    | 1    | 5,934,037 | 7924   | 1.58E+09    | 920000   | 678000  | 1020000  |
| LD    | 0.79 | 8,961,989 | 6831   | 759000000   | 817000   | 397000  | 1170000  |
| YH    | 1    | 5,502,967 | 600    | 2.00E+08    | 68300    | 50300   | 79700    |
| EM    | 0.78 | 4,835,928 | 490    | 245000000   | 59000    | 44900   | 70900    |
| SW    | 0.89 | 5,624,696 | 208    | 52000000    | 27600    | 17500   | 37300    |
|       |      |           |        |             |          |         |          |

Table 4. Estimated deaths number by regions, number of deaths per million, number of cases. 5% cases representthe least predicted number of cases and 95% cases represent the most.

## **IV.** Conclusion

The extreme complexity of events that occur due to a large number of uncertainties leads to the inherent deterministic unpredictability of these events, and the timing of vaccine interventions for vulnerable populations in different regions and ages can be well described using a Bayesian hierarchical prediction model. We developed an extended Bayesian hierarchical prediction model to forecast cases' numbers as well as the reproduction after vaccination in nine regions of England. Based on the population of each region, the number of deaths and the mortality rate from infection (IFR) in each region were predicted. We found that IFR was highest in the East, Northwest, and Southeast, corresponding to 29,079, 28,734, and 25,201 deaths, respectively. On January 7, 12, 16, 13, 10, 17, 10, 18 and 14, the regeneration number ( $R_t$ ) in each region will be reduced to less than 1 respectively. Given the infection fatality rate and epidemiological characteristics, major vaccine interventions have been effective in reducing transmission in nine areas in England. We also found that vaccination among people aged 70 to 80 has made a significant contribution to reducing transmission of the virus.

This model can be extended to predict the effectiveness of certain interventions in public health emergencies, the effectiveness of prevention of disease transmission, and the effectiveness of different interventions in different age groups, so as to find the best way to control the spread of disease. It can also be extended to pharmacological and non-pharmacological interventions to find the best solution.

## References

- [1] https://coronavirus.jhu.edu
- [2] https://baijiahao.baidu.com/s?id=1688641939538567669&wfr=spider&for=pc
- [3] T.,Lietman,,S.,M.,& Blower.(2000).Potential Impact of Tuberculosis Vaccines as Epidemic Control Agents.Clinical infectious diseases : an official publication of the Infectious Diseases Society of America,30(Supplement\_3),S316-S322.
- [4] Sheng-I,Chen,Chia-Yuan,& Wu.(2020).A stochastic programming model of vaccine preparation and administration for seasonal influenza interventions. Mathematical biosciences and engineering: MBE.,17(4),2984-2997.

- [5] Mohamed, El, Fatini, Idriss, Sekkak, Aziz, & Laaribi. (2019). A threshold of a delayed stochastic epidemic model with Crowly–Martin functional response and vaccination. Physica, A. Statistical mechanics and its applications, 520, 151-160.
- [6] Kuangang, Fan, Yan, Zhang, Shujing, Gao, Shihua, & Chen. (2020). A delayed vaccinated epidemic model with nonlinear incidence rate and Lévy jumps. Physica A, 544, 123379.
- [7] E. Mahase, Covid-19: UK records first death, as world's cases exceed 100 000. BMJ. 368 (2020), doi:10.1136/bmj.m943.
- [8] https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-10-march-2020.
- [9]https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-22-march-2020.
- [10] Seth Flaxman et al. "Report 13: Estimating the number of infections and the impact of nonpharmaceutical interventions on COVID-19 in 11 European countries". In: (2020). URL: https ://doi.org/10.25561/77731.
- [11] Michaela A. C. Vollmer et al. "Report 20: Using mobility to es mate the transmission intensity of COVID-19 in Italy: A subnational analysis with future scenarios". In: (2020).
- [12] https://coronavirus.jhu.edu/map.html
- [13] Li, R. et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science 368, 489–493 (2020).
- [14] Jombart, T. et al. Inferring the number of COVID-19 cases from recently reported deaths. Wellcome Open Research 5, 78 (2020).
- [15] Samy Dana et al. "Brazilian Modeling of COVID-19: a Bayesian Monte Carlo approach for COVID-19 spread in a limited data set context".
- [16]S,Eubank,I,Eckstrand et al. Commentary on Ferguson, et al., "Impact of Non-pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand". The bulletin of mathematical biophysics,82(4),1-7.(2020)..
- [17] Verity, R. et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet 20, 669–677 (2020).
- [18] Bi, Q. et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. Lancet Infect. Dis. https://doi.org/10.1016/S1473-3099(20)30287-5 (2020).
- [19] https://www.gov.uk/
- [20] TA Mellan, HH Hoeltgebaum, S Mishra et al. Estimating COVID-19 cases and reproduction number in Brazil. Imperial College London (08-05-2020), doi: https://doi.org/10.25561/78872.
- [21] Fraser, C. Estimating Individual and Household Reproduction Numbers in an Emerging Epidemic. PLOS ONE2, e758 (2007).
- [22] Cori, A., Ferguson, N. M., Fraser, C. & Cauchemez, S. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. Am. J. Epidemiol.178, 1505–1512 (2013).
- [23] Nouvellet, P. et al. A simple approach to measure transmissibility and forecast incidence. Epidemics22, 29–35 (2018).
- [24] Cauchemez, S., Valleron, A. J., Boëlle, P. Y., Flahault, A. & Ferguson, N. M. Estimating the impact of school closure on influenza transmission from Sentinel data. Nature452, 750–754 (2008)
- [25] Seth Flaxman, Swapnil Mishra, Axel Gandy et al. Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries. Imperial College London (30-03-2020) doi: https://doi.org/10.25561/77731.
- [26] Liu,Gayle,A,Wilder-Smith,& Rocklöv.(2020). The reproductive number of COVID-19 is higher compared to SARS coronavirus. Journal of Travel Medicine,27(2).
- [27] AB Hogan, P Winskill, OJ Watson et al. Modelling the allocation and impact of a COVID-19 vaccine. Imperial College London (25-09-2020), doi: https://doi.org/10.25561/82822.
- [28] https://www.google.com/covid19/mobility/