BRADFORD INSTITUTE FOR HEALTH RESEARCH
MAKING RESEARCH REAL

BIHR COVID-19 Scientific Advisory Group Working Paper

An overview of models for predicting demand for COVID-19 hospital inpatient care in local areas of the UK

Bo Hou[1]*, Jane West[1], Trevor Sheldon[2], Brian Kelly[1], Josie Dickerson[1], Richard Cookson[4], Mark Mon-Williams[3], Chris Cartwright[1], Kate Pickett[2], John Wright[1]

1 Bradford Institute for Health Research, Bradford, BD9 6RJ, UK
2 Department of Health Sciences, University of York, UK, York, YO10 5DD, UK
3 School of Psychology, University of Leeds, Leeds, LS2 9JT, UK
4Centre for Health Economics, University of York, York, YO10 5DD, UK
*Corresponding author: Dr Bo Hou, email address bo.hou@bthft.nhs.uk
Website www.bradfordresearch.nhs.uk/c-sag
ISBN 978-1-5272-6176-1

Abstract

Objective

We aimed to critically appraise the value for UK local health service planners of five major COVID-19 models (Imperial College, Sheffield, Edge Health, Oxford, and Draper & Dash) with respect to: the prediction of the peak timing and magnitude of infection, and the demand for hospitalisation.

Results

The Imperial College model is the most statistically advanced and robust model, useful at a national level. But it lacks transparency and adjustability at local level. A flexible and practical model, such as that from Sheffield, lacks sophistication but is potentially the most useful to local health services planners. It is easy to update with new parameter estimates related to both the transmission of the virus, and emerging evidence in a local area regarding clinical factors, so helping local health services monitor a dynamic situation. The latest projection from both the IC model and the results from the revised Sheffield model suggest the timing of peak infection and hospitable demand will be around mid-April 2020. Some local areas may experience later peaks due to differences in the onset of initial seeding of infection.

Keywords: Covid-19 models, hospital inpatient care, pandemic

Introduction

This paper was written in the first week of April 2020, in the early stage of the COVID-19 epidemic in the UK. Local health services in the UK need an idea of when hospital demand is likely to peak. The aim of this overview is to provide a practical assessment of different UK COVID-19 models from the perspective of planning local health service responses. We summarise and critique the main UK COVID-19 models, with particular relevance to Bradford, (a large metropolitan district in the North of England with high levels of deprivation, ill health and ethnic diversity). We looked at the models from Imperial College (IC), Oxford, Sheffield, Edge Health, and Draper & Dash. We hope this summary will be useful to other local areas.

The available models can be grouped into three broad methodological categories; a) individual-based discrete time simulation (IC model), b) epidemiological compartmental (Sheffield, Edge Health and Oxford), c) times-series econometrics forecasting (Draper and Dash) (see Table 1).

All models estimate health care demand using two steps. The first step involves estimating the number of infections in the population based on COVID-19 transmission characteristics, such as the reproduction number, the incubation period (time from infection to symptoms) and the latent period (time from infection to infectiousness). The second step is the estimation of demand for hospitalisation care, this is done based on assumptions around COVID-19 clinical characteristics such as the proportion of infected patients needing hospitalisation and the type of care needed (critical care or non-critical care).

Table 1. Summary of models reviewed

Model	Predicted Peak of Hospital Demand	Predicted Peak Number of all Hospital Beds	Strengths	Limitations	Local Adaptability
Imperial College	Mid-April 2020	5000 in ICU	Robust & advanced	Complex	Low
Sheffield (v0.6)	10th June (Bradford – mid April)	802 per 520,000 population	Practical & Flexible	Deterministic	Good
Edge Health	End May (no control measures) July-August (with control measures)	30,000	Straightforward	Crude	Good
Oxford	n/a	n/a	Highlights need to conduct serological survey	Strong assumptions	Possibly
Draper & Dash	n/a	n/a	Accessible, uses individual data	Forecasting based on data from other countries	Likely Good

The next section describes each of the models in more detail, focussing on the key issues in modelling the peak timing of infection and the associated demand for health care.

Models analysed

Imperial College (IC) model

The model in the 16th March report from Professor Neil Ferguson and colleagues is based on papers that modelled pandemic influenza. [1,2] The IC model is agent-based and, to our understanding, has been adjusted to predict the time course of covid-19 infection across the UK population. These types of models simulate systems containing autonomous interacting decision-making entities - 'agents' – allowing the impact of changing various parameters to be estimated. In order to create the simulation, the model uses demographic data – a virtual population comprising agents dispersed across geographical space. Household sizes, workplace sizes, age distributions, and proportions of children in school as a function of age are used as parameters for the demographic model. Data on the distances travelled to work are incorporated to estimate agent movement. The model also uses virus transmission characteristics - such as the distribution of the incubation period (the time from infection to symptoms). The disease parameters are based on data from known exposure events. Infectiousness is incorporated within the model using Bayesian methods with the assumption that infectiousness only starts at the end of the incubation period. The generation time (the average time between the infection of an individual and of their contacts) is modelled by considering the estimated latent period and the profile of infectiousness $k(T)$ with an assumed normal distribution. The generation time is modelled from the mean latent period plus $\int_0^\infty Tk(T)dT$ (this integral represents the average time after the latent period to infection of a contact). The simulations are stochastic and so provide margins of error for estimated values. These distributions include age and household sizes with local density determined by the dataset. The model simulates the system at discrete time steps. A given individual (agent i) has a probability of being infected, $1 - \exp(-\lambda i\ \Delta T)$, at each time step, where ΔT is the time step and λi is the infection risk of the individual. The infection risk is assumed to arise from three areas – household, place and random contacts in the community from movement.

In an interview on 27th March,[3] Professor Ferguson said that the only thing he would change now (based on the latest data from Europe) is Ro - from 2.4 (a value also used in many other models/reports) to slightly above 3. The implication is that more people in the population would already have been infected with the virus than originally estimated. This greater transmissibility plus the effect of total lockdown in the UK, has led to predictions of intensive care unit demand peaking in approximately 2.5 to 3 weeks' time (around mid-April 2020) and then declining. This is two months earlier than his original report.

Sheffield (version 0.6), Oxford and Edge Health model

These are essentially compartmental models used in epidemiology, which separate the total population into different compartments. The total population is equal to the total of these compartments at any time. Given the limited public availability of the IC model, Dr Chris Gibbons at Sheffield City Council developed the 'Sheffield' model and has shared this with Bradford Metropolitan District Council (among others). This model can be run in Excel and is regularly updated. In essence, it is a simple SEIR compartmental model that separates the total population into different compartments (Susceptible, Exposed, Infectious, Recovered) and uses parameter estimates from the IC report and Public Health England (PHE). The Edge Health model is based on a crude (cf complex) theoretical susceptible-infected-recovered (SIR) model with assumed model parameters[3]. It is crude in the sense that unlike the Sheffield model it did not use parameters from clinical studies on COVID-19. The Oxford model is also essentially a compartmental model, specifically, a susceptible-infected-recovered (SIR) model with a Bayesian MCMC approach.[4] A significant limitation of compartmental models is that they are deterministic and do not provide information about the potential error of estimates.

The Sheffield model is a good representation of a compartmental model and has been adjusted for local use. We focus on this model here. At each time point (day), the number of people in each compartment is calculated using the equations below:

The Susceptible population at period t:

$$S_t = S_{t-1} - \frac{Ro}{average\ infectious\ period} \times \frac{S_{t-1}}{N} \times I_{t-1} + birth\ rate - death\ rate$$

[3] https://www.edgehealth.co.uk/post/covid19

The Exposed population at period t:

$$E_t = E_{t-1} + \frac{Ro}{average\ infectious\ period} \times \frac{S_t}{N} \times I_{t-1} - \frac{1}{average\ latent\ period} \times E_{t-1}$$

The Infected population at period t:

$$I_t = I_{t-1} + \frac{E_{t-1}}{average\ latent\ period} - \frac{1}{average\ infectious\ period} \times I_{t-1}$$

The Recovered population at period t:

$$R_t = R_{t-1} + \frac{1}{average\ infectious\ period} \times I_{t-1}$$

At initial period (t=0), all of the population is assumed to be susceptible.

In the Sheffield model v0.6, the parameters used were: Ro=2.4, average latent period=5.1, average infectious period=4.6, whilst birth rate and mortality rate vary by age group. These were taken from the IC report and largely align with other reports. If the initial infection numbers and total population size is known then this model is suitable for local use. It also allows for age-specific prediction allowing adjustment for differences in local demography. This type of model demonstrates good flexibility in a local setting. Operation of other compartmental models would be similar to this model. To model demand for hospitalisation care, a set of assumptions were needed - for instance, the percentage of symptomatic cases requiring hospitalisation, percentage of infected cases requiring critical care, and infection fatality ratio. These were also taken from the IC report.

Draper and Dash model

Another category of empirical COVID-19 model uses time series forecasting. The Draper and Dash model (developed by an artificial intelligence company who have previously worked with NHS Trusts) is, to the best of our knowledge, a time series regression model. It probably used the Chinese and Italy data to estimate parameters and then a variant of an autoregressive conditional heteroskedasticity model (ARCH) to make forecasts. However, the quality of forecasting will be noisy as there will be few initial cases in any local area and forecasts are based on data from other countries.

Prediction of peak timing and magnitude of infection

As we wrote this, demand for hospital treatment was increasing rapidly (possibly doubling every 2-3 days) and services needed a prediction of the peak timing and the magnitude of infection and demand for hospitalisation. Using the IC model, Ferguson (27th March, 2020) projected that the peak in the UK would be mid-April 2020 rather than a previous estimate of around June 2020. The changes were due to higher estimates of virus transmissibility and the implementation of full lockdown in the UK.

One key parameter for these models is the timing of initial seeding in the population - when and how many people were first infected - as these are a function of time and changes each day. The initial seeding of infection for version 0.6 of the Sheffield model was 17th March 2020 and the peak of infection/hospitalisation was projected to be 10th June 2020. But it is very likely that the first infection was earlier but not reported. The Oxford model instead bases its projections on the timing of the first death, as this is likely to be better recorded. The interval between first infection and first death from European data is about one month. The first death in the UK recorded as being due to COVID-19 was 5th March 2020,[4] which suggests that the first infection in the UK population was around 5th February 2020. Both the IC and Oxford models used initial seeding of infection around the end of January. When we ran the Sheffield model using a more plausible date for seeding, we estimated a new projected peak in the last week of April, six weeks earlier. If we also use a Ro of 3 rather than 2.4, the peak would be around two weeks sooner (around mid-April). This is because more people would have been infected, recovered and be immune. The only way to test whether a sizeable proportion of the UK population may already have been infected is to conduct a large serological survey. Furthermore, increasing Ro also increases the size of the peak hospitalisation demand (other

[4] https://en.wikipedia.org/wiki/2020_coronavirus_pandemic_in_the_United_Kingdom

variables remaining constant) as more people would get infected in the same period of time relative to a lower Ro.

Assumptions made about the effectiveness of the lockdown in reducing the transmission of COVID-19 (measured by Ro) are also important in predicting the peak timing and size of the epidemic. Ideally, we want to reduce Ro to less than 1. We demonstrated this in two scenarios, see the Appendix S1.

There are important uncertainties related to the pathology of the virus and unobserved variables in all of these models. Model projections should ideally be adapted to account for variations in population density, socio-demographics, cultural and other factors influencing behaviours and rates of underlying morbidity. All the models have assumed the characteristics of COVID-19 are stable and don't, for example, vary over time and with changes in the weather. The projection of health care demand not only depends on the transmission mechanism of the virus, but also on the clinical parameters related to the treatment of the virus, such as hospitalisation rate, recovery time and length of hospital stay.

Conclusion

We have assessed current major COVID-19 models in the UK. The IC model is the most statistically advanced and robust model. But it lacks transparency and adjustability at a local level. A simple, flexible and practical model, such as that from Sheffield, lacks sophistication but is potentially the most useful to local health services planners. It is easy to update with new parameter estimates related to both the transmission of the virus, and emerging evidence in a local area regarding clinical factors, so helping local health services monitor a dynamic situation. Underlying health service supply capacity is different across local areas which also supports the need for a practical and flexible local model. This type of model may also inform local mental health and social services planning, as demand for other services is likely to be positively related to the number of infections. The practical nature of this type of model may also facilitate communication across different agencies in local area.

The latest projection from both the IC model applied to the national population, and the revised Sheffield model applied to a population like Bradford, suggest the timing of peak infection and hospitable demand in the UK will be around mid-April 2020. It is likely there will be regional differences in peak time of infection/demand due to differences in the timing of initial seeding of infection. The magnitude of the peak is important as local health resources, such as critical care hospital beds, could be quickly overwhelmed by a rapid increase in demand.

Limitations

- We have not examined all the COVID-19 models in the UK.
- As the methodology of some of these models are not publicly available, some of our work is based on best knowledge.

Availability of data and material

No data are associated with this article.

Acknowledgement

We want to thank all frontline health care workers around the world who are fighting COVID-19. We thank our colleagues at Bradford District COVID-19 Scientific Advisory Group and our wonderful colleagues at Bradford Institute for Health Research Scientific Advisory Group for their support in developing this document. Also, we would like to thank Ruaridh Mon-Williams and Kirsty Mon-Williams for their helpful inputs.

Author's contribution

BH wrote this article. JW, TS, BK, JD, MMW, RC, CC, KP, JW supervised, edited and reviewed this article.

Funding

Not applicable.

Ethics approval and consent to participate

No ethical approval was required for this project as this is secondary research.

Consent to publish

Not applicable.

Competing interests

Nothing to declare

References:

1. Ferguson N, Cummings D, Cauchemez S, et al. Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature
2005.
2. Ferguson N, Laydon., Nedjati-Gilani., et al. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID- 19 mortality and healthcare demand. Online
Imperial College COVID-19 Response Team, 2020.
3. Ferguson N. Let’s flatten the coronavirus confusion curve. In: Kelly J, editor.: Financial Times; 2020.
4. Lourenço., Paton., Ghafari., et al. Fundamental principles of epidemic spread highlight the immediate need for large-scale serological surveys to assess the stage of the SARS-CoV-2 epidemic. 2020.

Appendix

Assumptions made about the effectiveness of the lockdown

To illustrate this point, we used an online and public SEIR model tool[5] with parameters taken from WHO reports on COVID-19. We assumed a population of around 500,000 (approximate population of Bradford), Ro = 3.04, lockdown/ intervention date March 16th 2020 (10 days after first death). We considered two scenarios: first, where the intervention reduces transmission by 67% (to Ro=1) and second where it reduces it by 80% (to Ro=0.61). In scenario 1, peak hospitalisation timing is 112 days (end of May 2020) with demand of 1,882 beds. In scenario 2 peak hospitalisation timing is 67 days (mid-April 2020) with demand of 846 beds. This highlights the importance of bringing the transmission rate as low as possible in our communities.

[5] http://gabgoh.github.io/COVID/index.html