Written evidence submitted by Professor Neil M. Ferguson , Imperial College London

(C190038)

Re: Commons Science & Technology Select Committee Inquiry "UK Science, Research and Technology Capability and Influence in Global Disease Outbreaks"

I write to respond to the written questions posed by the committee ahead of the oral evidence session I will be participating in on June 10th this year.

The MRC Centre I head has published 25 scientific reports on the COVID-19 epidemic between 17th January and 25th May this year. Over 50 staff and research students have been involved in the research undertaken. The code and data used to prepare these reports has been released in over 20 public *github* repositories (https://github.com/mrc-ide). While code is often reused across different studies, some customisation is nearly always required, so in a sense we have used as many models as we have published studies. Indeed, some studies use multiple models of different types.

Rather than list every model and every study, it is therefore perhaps more helpful to summarise some important aspects of infectious disease modelling in relation to COVID-19. In doing so, I hope I have addressed the technical questions posed by the Select Committee. The reports referred to below (most of which have been or are in the process of being published as peer-reviewed scientific papers) are available at http://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/covid-19-reports/.

1. Priorities for epidemiological research in emerging infectious disease outbreaks

Upon the discovery of an outbreak of a novel pathogen the initial priorities are to:

- Understand the scale of the outbreak, given initially outbreaks are nearly always under-reported (Reports 1, 2, 6 and 8).
- Quantify the extent and rate of human-to-human transmission, by estimating the basic reproduction number, R_0, the number of secondary cases per case (Reports 3, 13, 20, 21, 23).
- Quantify the risk posed to human health, principally by estimating the Infection Fatality Ratio (IFR), the proportion of those infected expected to die as a result and assessing likely overall population health impact (Reports 4, 8, 9, 12, 17).
- Assess the effectiveness of control measures and the implications for future policy formulation (Reports 9, 11, 12, 13, 16, 18, 20, 23, 25).

2. Data and scientific uncertainty during an emerging infectious disease outbreak

Scientific data is inevitably limited at the earliest stages of a novel infectious disease outbreak. One of the benefits of modern rigorous epidemiological analysis is to make optimal use of that limited data to derive maximum insight into the unfolding epidemic. Thus most of the research of the MRC Centre during this

outbreak has been the analysis of data – often using sophisticated statistical models to account for data gaps and uncertainty and to integrate data from a variety of sources to address key questions.

As an epidemic progresses, the available data grows (for instance, due to improvements in surveillance), allowing more precise estimates of key variables (e.g. R_0 and IFR) to be made, and more complex questions to be addressed (e.g. understanding risk factors for severe disease). Sometimes, new scientific discoveries fundamentally change our understanding of a disease; however that is not the case for COVID-19 – the conclusions we had drawn by the end of February regarding transmissibility and severity have been reinforced and confirmed by later data.

We have made use the following range of data sources through the epidemic:

- Confirmed case counts. Often the most widely available data, but subject to significant biases over time and between countries due to variation in the availability of and criteria for testing
- Death data. Still subject to some surveillance biases and therefore under-reporting, especially early in an epidemic, but typically more reliable than case data, since surveillance systems are typically more tuned to detecting severe disease.
- Infection prevalence data – obtained for COVID-19 by swabbing everyone in a group (e.g. plane, cruise ship, hospital) and testing the samples by PCR for presence of the virus. These data were critical to making early assessments of the severity (IFR) of the virus.
- Seroprevalence data – obtained via antibody testing. These are typically available later in an epidemic. Such data are invaluable for refining estimates of severity and for assessing the extent of transmission in a population if many infections are mild or asymptomatic (as for COVID-19). Seroprevalence data requires careful analysis to correct for limitations of the tests used (i.e. imperfect sensitivity and specificity).
- Contact tracing data – this gives insight into who is more likely to infect who, for instance allowing the incubation period and serial interval distributions for a new disease to be characterised (two key quantities), and giving insight into the risk of transmission in different settings (e.g. households or workplaces).
- Health system utilisation data (e.g. new hospital admissions, ICU beds occupied) – this allows better characterisation and forecasting of health system impacts of an emerging epidemic. Modelling healthcare demand requires additional model parameters characterising the probability of hospitalisation and the average length of stay; for the UK, these have been derived from the data collected by the ICNARC consortium (https://www.icnarc.org/).
- Pathogen genomic data – obtained via full-genome sequencing of viral samples. This allows evolution of the virus to be tracked, but also gives invaluable epidemiological data – e.g. on seeding of infection from one area of the world to another, or on the rate of growth of the epidemic.
- Demographic data – this is used by all mechanistic transmission models to accurately capture the structure of the human population being modelled. Simpler models only require data on the age distribution of a population, while more complex models may stratify by region or risk factor – or capture more detailed aspects of population structure such as households, schools and workplaces.
- Contact survey data – data on the frequency with which people make different types (e.g. physical, conversational) of contacts with each other is invaluable in modelling respiratory viruses. These data are typically collected through surveys or contact diary studies. Nearly all the models used by SPI-M make use of such data to parameterise, for instance, the frequency with which someone in one age group contacts people of other ages. Furthermore, weekly surveys conducted by LSHTM through the COVID-19 pandemic have given insight into how contact patterns have changes under lockdown. The one caveat for these data is that the contacts measured are a (likely imperfect) proxy for infectious contacts.
- Mobility data - a new source of data which has risen to prominence in this epidemic has been data on population mobility, collected via mobile phones. This has given quantitative measures of the impact of social distancing on travel and mobility (and therefore likely contact patterns) in different settings.

It has been integrated into two of the MRC Centre's models, and several more in SPI-M.

No single data source gives a complete picture. We therefore focus on analyses which integrate multiple data streams to address the priority questions listed above.

3. Types of and relationships between models

Perhaps the key distinction in types of models used in epidemiology is between non-mechanistic and mechanistic (also called dynamical) models:

- *Non-mechanistic*: these are purely statistical models (e.g. regression or other forms of distribution- or curve-fitting) intended to estimate epidemiological parameters (e.g. IFR, incubation period, serial interval).
- *Mechanistic (dynamic)*: these embed knowledge about the biology and epidemiology of a virus and represent aspects of the transmission process. The simplest is a 'branching process' model (essentially an exponential growth assumption), while more complex models generally fall into the "SIR" category – modelling individuals flowing from being initially Susceptible to infection, to being Infected, to being Recovered (or dead). These models all capture the transmission dynamics of an epidemic.

Both classes of models can be fitted to data, thanks to modern statistical methods, though for mechanistic models there is a trade-off between complexity and the ease (computationally and statistically) with which they can be directly fitted to data or need to be parameterised in other ways (e.g. by using parameter estimates derived from simpler models).

Mechanistic SIR models can be divided into two classes: compartmental models, which track the numbers of people in different states (e.g. susceptible, infected or immune) and individual-based models, which simulate every individual in a population. Compartmental models have a long history: the models being used by us, Warwick and LSHTM were first developed 30 years ago, though all three groups have recoded their implementations of that basic model structure from scratch for this epidemic. Individual-based models are a little more recent and less standardised, but the mathematical design of the simulation we use (representing households, schools and workplaces) was first published >20 years ago.

In the MRC Centre we use both, though most of our work uses simpler compartmental models, due to their lower computational requirements and relative simplicity. An exception is the simulation used for Report 9 and some of the ongoing work for SPI-M. We chose to use a more complex model in that context due to its ability to represent household based interventions and also because by doing so, we provided a greater diversity of model types within the work being undertaken for SPI-M and SAGE (given the main LSHTM and Warwick models were both of the age structured compartmental SIR type).

While individual-based models are more computationally complex (to both program and run), they do not necessarily make more assumptions or require more parameters than compartmental models. The age-structured SIR compartmental models used by us, LSHTM and Warwick each have ~150 transmission parameters (elements of a matrix determining the risk that someone of one age will infect someone of another age). These are all estimated from contact survey data, but the applicability of those surveys to COVID-19 is an assumption. Conversely, the individual-based model used for Report 9 has around 10 transmission-related parameters – again mostly derived from contact survey data. The smaller number of transmission parameters is because the more detailed representation of the population (which requires much more demographic data to represent) in individual-based simulations intrinsically constrains potential transmission pathways.

Generally, we use the simplest model possible to address a specific question, since the more complex the model, the more parameters need to be estimated or assumed. There is little benefit in using a complex simulation to estimate R_0, but conversely, if one wants to assess, for instance, the likely effect of household quarantine on disease transmission, it is necessary to use a simulation which includes households (which most simpler SIR models don't).

A further distinction between mechanistic models is whether they model transmission as a random process (stochastic models, implemented as Monte Carlo simulations) or just characterise the 'average'

behaviour of an epidemic (deterministic models). Within the MRC Centre, we use both classes of models. Deterministic models are much less computationally intensive to fit to data, while stochastic models can better characterise the random variation in the transmission process. Individual-based models are always stochastic.

As there are limits to how many parameters can be estimated at the same time in fitting a mechanistic transmission model to limited data, often only a subset of parameter are fitted, with others being fixed at values which have been estimated in previous analyses, using other models. Examples of these include the incubation period, contact rates between different age groups in the population, or the proportion of transmission occurring in households versus schools.

4. Uses and outputs of mechanistic models

The uses of mechanistic models also vary, there being two broad classes of applications:

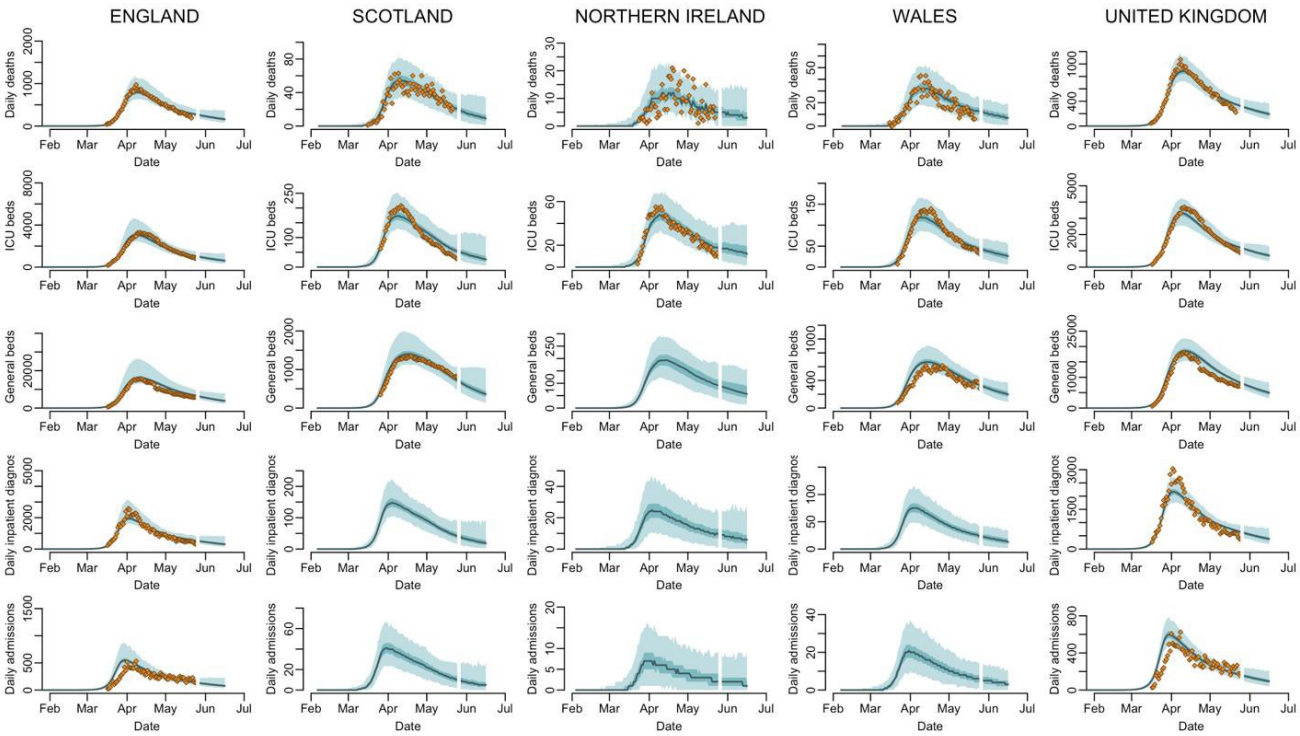
- *Scenario modelling*: using models to project possible future long-term trajectories of an epidemic, and the potential impact of different control strategies (e.g. Reports 9 and 12). When only very limited surveillance data are available in a country (as was the case in the UK in late February and early March), such scenario modelling takes parameter estimates drawn from analysis of data in other countries (e.g. China) and extrapolates to the UK. As local data become available, it becomes possible to fit models to the trajectory of the epidemic so far, and then project forward the possible impacts of, for instances, changes in control policies. For a new virus, assumptions inevitably need to be made about the likely effectiveness of control measures (e.g. the level of population compliance with social distancing, or the proportion of cases identified via contact tracing), so such scenario modelling cannot make statistically rigorous predictions. More, it is a useful tool to inform policymakers of the potential impacts of different policy choices under a range of assumptions for the effectiveness of different policies.

  Hence assumptions (e.g. around adherence to social distancing) made at one point in time are later refined when data become available on how the population responded to a policy which was implemented. Generally, UK population adherence to the policies modelled in Report 9 have been as good or better than we assumed in that analysis. Report 9 did not attempt to make precise predictions of the future mortality from COVID in the UK under different intervention options since at the time the research was conducted it was impossible to reliably assess how far the epidemic had progressed, plus final policy decisions not been made. Rather we used sensitivity analysis to examine a wide range of scenarios for intervention timing and policy choices, showing how outcomes varied with those variables and with the assumed transmissibility of the virus (R_0).

- *Real-time forecasting*: using models to make short-term predictions of the trajectory of an epidemic (e.g. Report 13 and https://mrc-ide.github.io/covid19-short-term-forecasts/index.html). Here the focus is on fitting models rigorously to multiple data streams (e.g. deaths, hospital admissions, ICU beds occupied, serology) and predicting the course of the epidemic over the next few weeks, with uncertainty fully characterized. The outputs are probabilistic – a distribution of possible trajectories, representing both statistical uncertainty in model parameters and, for our models, stochastic variability. Often the outputs from multiple models are combined in making such forecasts, since it has been established that such "ensemble" forecasts perform better than any one individual model. This is the process that SPI-M adopts to generate weekly estimates of "R" and infection incidence.

  In terms of outputs, the simplest models we use focus only on estimating aggregate deaths and infections over time (e.g. Reports 13, 20, 21, and 23), while more complex models (e.g. Reports 9 and 12 and the real-time modelling undertaken for SPI-M) also output estimates of healthcare demand (e.g. hospital admissions and ICU occupancy), sometimes stratified by age group and/or region. Parameters required for the modelling of healthcare demand are derived from detailed patient data (e.g. for the UK, collected by the ICNARC consortium - https://www.icnarc.org/). For examples of outputs, we refer the committee to our published reports at http://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-

19/covid-19-reports/, but the figure below gives an example of a recent real-time modelling fit to UK data:



5. Software implementation of models

We use a variety of programming languages to implement models, with the most frequently used being the statistical language 'R' and C/C++. R is a high-level interpreted language well suited to scientific computation, and for which a wide variety of open-source packages are available for statistical and epidemiological modelling. However, it is relatively slow, so computationally intensive code is still typically written in C or C++. Often both languages are combined (R can call code written in C/C++). Models are run on a variety of operating systems; most individual researchers use Windows or MacOS on their desktops, while we have access to both Windows and Linux high performance computing clusters.

Over the last 7 years, the MRC Centre has invested in developing an internal professional research software engineering team to improve the standard of internal software development and to adopt best practice in software reuse and research reproducibility. Where code is likely to be used in many future projects (by ourselves or others), we invest additional resource and effort in design, documentation, and testing. However, no epidemiological research group has the resources available to commercial software companies.

Since SIR transmission dynamics are common to a wide range of diseases, the models used for COVID closely resemble and are sometimes based on those used to model influenza, measles and other viral pathogens. However, all the COVID models we have used have been specifically parameterised for COVID (for instance, in specifying R_0, severity, the incubation period and serial interval).

6. Modelling as a scientific input to policymaking: model validation and comparison

The UK has a sophisticated and well-established system for ensuring that policymaking in infectious disease crises is informed by science, first established in the BSE and foot-and-mouth epidemics, then later refined during SARS, the 2009 flu pandemic and Ebola. Epidemiological modelling has played an important role in all of those epidemics, and has a dedicated ad-hoc group, SPI-M. SPI-M exists to marshal the wide range of epidemiological modelling expertise in the UK (in both universities and PHE). SAGE and SPI-M are not passive recipients of modelling; most of the work undertaken by academic groups contributing to SPI-M is commissioned by SAGE. SPI-M works on the principle that no single group has

uniquely perfect insight, and therefore that no policy decision should depend on a single group or model.

Real-time research in an epidemic inevitably occurs on a much faster timescale than scientific understanding progresses at other times, and nearly always in a context of very limited data and therefore substantial scientific and statistical uncertainty. The pace of research in this pandemic has been truly exceptional, but one consequence of this is that much of the work informing both new research and scientific advice is in the form of "pre-prints" rather than peer-reviewed papers (which lag weeks to month behind pre-prints). Furthermore, peer-review is a very imperfect process – for instance, it is very rare for code to be reviewed as part of the peer review process. Thus while the core code used in Report 9 underpinned two peer reviewed papers in the top scientific journal Nature in 2006 and 2007, it is only as a result of the Royal Society's current RAMP initiative and independent initiatives such as Codecheck (used by both ourselves and LSHTM - https://codecheck.org.uk/) that it is being more formally assessed.

For those reasons, SPI-M relies on comparing epidemiological estimates and modelling results produced by independent and in normal circumstances, competing academic groups to produce consensus statements which inform SAGE deliberations and then, potentially, policy.

During this epidemic, both we (Report 4 and https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30243-7/fulltext) and LSHTM derived independent estimates of the severity (IFR) of COVID-19, using different data sources. We estimated an IFR of 0.66% (95% CI: 0.39%-1.33%) for China and estimated that would translate to 0.9% in the UK, given our older population. LSHTM derived broadly similar estimates (https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30243-7/fulltext). These estimates informed the 'reasonable worst-case' planning scenario for COVID-19 adopted by SAGE in March. We would note that our current estimates of the IFR for the UK (combining mortality and serology data) are in the range 0.7-1.1%.

Multiple SPI-M groups also concluded by early March that R_0 for COVID-19 was likely in the range 2-3 (current best estimates for the UK are in the 2.8-3.2 range, though considerably uncertainty remains), and thus that 60-80% of the UK population might become infected if no control measures were adopted.

It was because of these conclusions about severity and transmissibility that multiple groups – ourselves and (our Report 9) and LSHTM (see https://cmmid.github.io/topics/covid19/reports/uk_scenario_modelling_preprint_2020_04_01.pdf) reached similar conclusions in early March that intensive non-pharmaceutical interventions would need to be implemented in the UK to avoid the NHS being overwhelmed.

Within the MRC Centre we also use a range of models to test the robustness of key results. We used a much similar age-structured deterministic SIR model (similar to that used by LSHTM) to both replicate the conclusions of Report 9 and then extend that work to consider the impact of COVID-19 in nearly every country of the world while accounting for health system and demographic differences (Report 12, peer-reviewed paper in press). Furthermore, we have verified that by switching off different features of our individual based simulation, we can exactly reproduce the dynamics of simpler models.

Last, while SPI-M provides inputs to SAGE via its consensus statements, these statements highlight both consensus and uncertainty. It is rare for different models to produce identical results, both due to differences in parameter estimates and in model design. Thus, it is more important to identify when the conclusions of modelling are qualitatively similar or different from a policy perspective. In relation to the results which informed lockdown there was qualitative (and a high degree of quantitative) agreement. However, in other areas (e.g. the potential impacts of reopening schools, or of the introduction of contact tracing and testing) there is less consensus, in large part because of continuing underlying scientific uncertainty (e.g. around the role of children in transmission or the infectiousness of those with asymptomatic infection).

Yours sincerely,

Neil Ferguson