Witness Name: Neil Ferguson

Statement No. :1

Exhibits: 384

Dated: 11 July 2023

UK COVID-19 INQUIRY

MODULE 2

WITNESS STATEMENT OF PROFESSOR NEIL FERGUSON

---

1. I, Professor Neil Ferguson, Imperial College London, Exhibition Rd, South Kensington, London SW7 2BX , will say as follows:

Table of Contents

A. Introduction

2. Given the scope of this submission, I have structured it into sections. I initially provide background on my role, the research units I direct at Imperial College London, and on epidemiological modelling. Later sections then cover the work undertaken by the "Imperial College COVID-19 Response Team" (ICCRT) I led at Imperial to inform the UK response to the COVID-19 pandemic. I conclude with some reflections and suggestions for change relating to how the UK plans and responds to civil contingencies (focussing on infectious disease threats) and the role of science and modelling in informing policymaking in crises.

3. Appendix A provides a chronology of the pandemic, from largely a UK perspective. It highlights the progression of the epidemic, significant scientific advances (with a focus on the work of the ICCRT and policy measures introduced in England).

4. In the text below, I have focussed on the research by ICCRT which was most relevant to the UK response to COVID-19. I provide a detailed but not exhaustive summary of this work. I have not detailed the large volume of research we undertook which focussed on supporting international organisations (notably the World Health Organization, WHO) and countries outside the UK. However, all publications resulting from such work are listed in the comprehensive set of COVID-19 related research outputs I provide in Appendix B.

5. Superscripted numbers in the text are citations to documents listed in the bibliography (section Q). These include a number of email exchanges with government officials which I submit with this statement as exhibits. Where mentioned, citations are provided to SAGE documents and minutes; as these are in the public domain, I only provide URLs to those documents. While not in the public domain, I assume that the Inquiry has access to all SPI-M-O meeting documents and so have generally only referenced the dates of those SPI-M-O meetings rather than providing documents themselves.

B. My role

6. I am a mathematical epidemiologist employed by Imperial College London, where I have been a Professor since 2001. My research focuses on using statistical and mathematical models to understand infectious disease dynamics and control. I have worked on emerging infectious disease outbreaks since 1995. Much of this work has involved national and international policymakers.

7. I hold the following research-focussed positions at Imperial College:

   a. Director of the MRC Centre for Global Infectious Disease Analysis (MRC GIDA): 2007-

   b. Head of the Department of Infectious Disease Epidemiology, School of Public Health: 2012-

   c. Director of the Health Protection Research Unit (HPRU) for Modelling and Health Economics: 2014-

   d. Director of the Abdul Latif Jameel Institute for Disease and Emergency Analytics (Jameel Institute): 2019-

   The MRC Centre, NIHR HPRU and Jameel Institute are discussed below.

8. I am an author on over 200 peer-reviewed scientific papers and am an elected fellow of the UK Academy of Medical Sciences and an elected international member of the US National Academy of Medicine.

9. I served on the predecessor of SAGE during the 2001 foot and mouth (FMD) epidemic and on every infectious disease related SAGE since (pandemic influenza 2009, Ebola 2014, Zika 2016, COVID-19 2020).

10. During the COVID-19 pandemic, I participated in the following UK government advisory groups: SAGE, SPI-M-O, NERVTAG (a statutory advisory committee which acted as a SAGE subgroup during the pandemic), SAGE-EMG (Environmental modelling group) transmission subgroup , the Advisory Committee for Dangerous Pathogens (ACDP, also a statutory advisory committee, but not a SAGE subgroup), the PHE/UKHSA Variant Technical Group, plus a number of other ad-hoc SAGE task-and-finish (*e.g.* on the role of children in transmission) and PHE/UKHSA groups (*e.g.* on vaccine effectiveness, and on the effectiveness of Test and Trace). In addition, I presented research findings to a number of meetings of the Joint Committee on

Vaccination and Immunisation (JCVI) in 2021. I also attended a small number of informal small group discussion meetings involving Patrick Vallance and/or Chris Whitty throughout 2020 and 2021. My formal memberships of NERVTAG and ACDP were extended in 2022.

11. At no point during the pandemic did I have any meetings or discussions (written or oral) with UK government ministers. In the first few months of the pandemic, I and the team at Imperial had a number of direct interactions with civil servants in several government departments and with officials in the NHS. However, as the pandemic progressed, such interactions (other than those with PHE/UKHSA) were increasingly directed through the SPI-M-O secretariat to ensure consistency in the modelling outputs being used across government.

12. As head of a large epidemiological research centre, my contribution to the groups and committees I participated in was as much as a representative of ICCRT as an individual scientist, in that much of my contribution was reporting or discussing the results of the research of ICCRT. A large proportion of the that work was formally or informally commissioned or requested by those groups or government departments. While this was the standard modus operandi for most participants in SPI-M-O and the UKHSA Variant Technical Advisory group, it was not for most other committees or groups (e.g. SAGE or NERVTAG), where most members act solely in an individual capacity. In addition, I also provided more general scientific advice on COVID-19 epidemiology and transmission dynamics, and such general advice represented the bulk of my contribution to the EMG subgroup and ACDP.

13. I believe that scientists have a key role to play in advising policymakers on the potential impacts of different policy choices in a crisis, but that they should not use the public platform offered to them by that role to campaign or advocate for specific policies. This belief is informed by two considerations: (a) legitimacy – scientists may have more expertise in a topic but have no more democratic right to determine policy than any other citizen; (b) trust – it can be difficult to generate the necessary levels of trust between government and independent scientific advisors if those advisors are both advising privately on the policy implications of developing science and advocating (publicly or privately) for specific policy options. I appreciate other scientists hold different views, and that there is an argument that when decisions are time-critical, scientists may feel an overriding moral imperative to communicate their views on the necessity for specific actions. I comment further on public communication of modelling-based research in section N below.

C. MRC GIDA and the Imperial College COVID-19 response team

14. I founded the MRC Centre for Outbreak Analysis and Modelling in 2007, with financial support from the Medical Research Council and Imperial College London. Funding was renewed in 2012, and then again in 2017 when the name of the Centre was changed to the MRC Centre for Global Infectious Disease Analysis (MRC GIDA).

15. At its founding, the MRC Centre consisted of approximately 7 tenured academic staff (*i.e.* lecturer and above) and some 50 people total (academic, research and professional staff plus PhD students). By 2020, the Centre had grown to 32 academic staff and over 150 staff and PhD students in total.

16. The Centre funding provided by MRC represents under 10% of our total research funding. Major funders include UKRI (project grants from multiple research councils), NIHR, the Wellcome Trust, the Bill and Melinda Gates foundation, the US National Institutes of Health, and Community Jameel (a philanthropy based in the Middle-East). The Director controls core Centre funding, but not the other research grants held by other principal investigators (i.e. the academic staff) within the Centre.

17. As Centre Director, I have been supported by four Associate Directors since 2017: Profs Christl Donnelly, Azra Ghani, Timothy (Tim) Hallett and Nicholas (Nick) Grassly.

18. In practice, and in common with most academic centres, decision-making within the Centre is typically collective and based on consensus. In particular, the Director does not have the authority to tell academic staff what to work on. Rather, the role is more focussed on strategically-drive coordination of activity and supporting the research (and career progression) of staff within the Centre.

19. The figure below illustrates the overall structure of the Centre. Since 2017, MRC GIDA has had 5 major research themes: (a) Outbreak Analysis and Modelling; (b) Global Health Analytics; (c) Vaccines; (d) Antimicrobial Resistance; (e) Methods and Tools. It conducts research on nearly all major human infectious diseases, most notably "emerging" infections (including COVID-19) and the high health burden globally endemic diseases of HIV, TB and malaria. Work spans epidemiological and genetic analysis together with statistical and mathematical modelling. We also conduct some field (epidemiological studies and clinical trials) and laboratory-based research. Most of our research is focussed outside the UK, principally

involving low- and middle-income countries where the majority of the health burden from infectious diseases is now experienced.

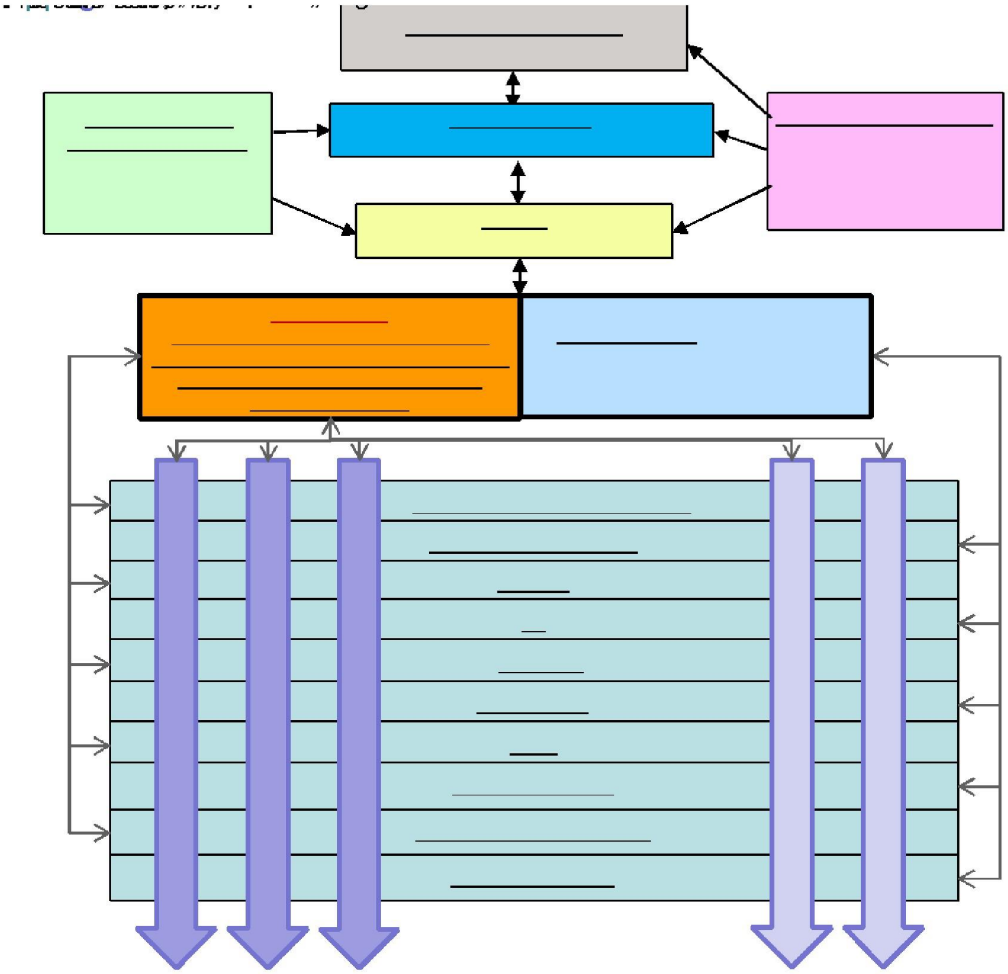

Figure: Structure of MRC GIDA. PI=Principal Investigator (*i.e.* academic staff of the Centre).

20. The mission of MRC GIDA is "*to be an international resource and centre of excellence for research and capacity building for the epidemiological analysis and modelling of infectious diseases*" and "*to undertake applied collaborative work with national and international agencies to support policy planning and response operations against infectious disease threats*".

21. Given our mission, we have established close working relationships with national and international public health bodies over the last 15 years:

    a. UKHSA (previously PHE): building on previous research collaborations, in 2014 Imperial College and PHE were awarded an NIHR grant to create the NIHR Health

Protection Research Unit (HPRU) for Modelling Methodology, a partnership between PHE and Imperial College London. In 2019 we were successful in a competitive bid to renew HPRU funding (this time for a unit titled the HPRU for Modelling and Health Economics), extending the HPRU partnership to include the London School of Hygiene and Tropical Medicine (LSHTM). The HPRU funds long term collaborative research between the partners and also provides "responsive mode" funding for short-term priorities. I have directed the HPRU since it was founded in 2014. In addition, I and other ICCRT staff contribute to two other NIHR HPRUs (again partnerships with UKHSA) hosted by Imperial College – in Respiratory Infections and in Healthcare Acquired Infections and Antimicrobial Resistance. We also have a number of staff with joint appointments between UKHSA and ICCRT (most notably Peter White, head of the modelling and economics unit at Colindale, and Steven Riley, now Director General for Data, Analytics and Surveillance for UKHSA), staff on long-term part-time secondments (Erik Volz), and a number of staff with honorary appointments at UKHSA.

b. The World Health Organization (WHO): in 2008, MRC GIDA was awarded WHO Collaborating Centre (WHO-CC) status as the first WHO-CC for Infectious Disease Modelling, building on the extensive links Centre staff had undertaken with WHO prior to the formation of the MRC Centre (e.g. on H5N1 avian influenza, SARS-1). WHO-CC status offers no funding but formalises a joint workplan across multiple areas of engagement and specifies the contexts under which WHO can call upon a CC for support. Since the formation of the CC, we have dramatically expanded our engagement with WHO, spanning not just work on epidemics and emerging infections, but also all major endemic disease threats (e.g. malaria, HIV, TB, dengue, yellow fever, hepatitis, polio, helminthic infections). In addition to work on COVID-19, in the past we have provided substantial analytical support to WHO during the 2009 H1N1 influenza pandemic (when I was an advisor on the WHO IHR Emergency Committee), during multiple Ebola outbreaks, and during the Zika epidemic in Latin America. Nearly all academic staff in MRC GIDA work with WHO, in many cases sitting on advisory panels.

c. Other international bodies: Over the last 15 years, we have developed close relationships with and have provided analytical support for most major international

public health organisations working on infectious diseases, including Gavi, the Global Fund, CEPI, PATH, the Bill and Melinda Gates Foundation (BMGF), the Wellcome Trust and the World Bank. Several of our staff and students have moved on to roles in those organisations.

d. Countries outside the UK: MRC GIDA researchers have research collaborations in approximately 80 countries world-wide, many low- or middle-income. Many of these links involve policymakers in those countries. Countries where there are particularly strong links (*i.e.* over many years, in multiple disease areas, or on largescale projects) include the USA, Singapore, Brazil, India, Zimbabwe, Malawi, Senegal, Colombia, Burkina Faso, Zambia and Indonesia.

22. The Jameel Institute at Imperial College was founded in 2019 with a substantial philanthropic donation (the majority used for capital investment in the new Imperial School of Public Health building) from Community Jameel, a large Middle Eastern philanthropy. Its remit is to use data analytics to address health emergencies, with a broader remit than MRC GIDA (*i.e.* including non-infectious disease health threats). The staff of the Jameel Institute overlap with those of MRC GIDA, and the work of the two bodies on the pandemic was entirely integrated. Hence much of the research on COVID-19 attributed to MRC GIDA in this document also benefitted from the financial support of the Jameel Institute and can also be attributed to the Institute.

23. MRC GIDA started working on COVID-19 on January 13th 2020. As the pandemic progressed, the collaborative team working on COVID-related research expanded beyond the MRC Centre (*e.g.* to include some staff from the Department of Mathematics at Imperial College), leading to our adoption of the name "the Imperial College COVID-19 Response Team" (ICCRT) in external communications from our Report 9 [NF/001 - INQ000049647] onwards. However, governance of the team made use of MRC GIDA structures.

24. I directed the work of MRC GIDA and ICCRT during the pandemic. From January to mid-March, work was organised via daily team meetings which I chaired. However, as the scope and scale of research expanded, my leadership shifted away from the previous detailed direction of (and participation in) most of the research undertaken (up to Report 12 [NF/002 - INQ000262593]), towards higher-level organisation and coordination of a set of otherwise largely autonomous streams of research:

a. UK Real-time modelling (RTM): responsible for most inputs into SPI-M-O, led by Marc Baguelin and Anne Cori. I retained a high level of often technical involvement in and strategic leadership of this stream throughout, guiding the work undertaken.

b. Global (non-UK) modelling: responsible for most inputs into global partners (WHO, Gavi etc) and for bilateral support of low- and middle-income country partners. Strategic and technical leadership was provided by Azra Ghani, with lighter-touch input from me (via management group meetings) than for the UK RTM stream.

c. Health economics and health systems: work initially developed models and software to assist in healthcare capacity planning, and then later focussed on developing integrated epidemiological and economic models capable of modelling both the economic and health impacts of control measures. This stream had a largely non-UK focus and was led by Katharina Hauck, with light touch input from me.

d. Spatiotemporal modelling of the pandemic, including estimating the impacts of interventions: this research was led by Samir Bhatt in MRC GIDA, but with substantial input from colleagues in the Dept. of Mathematics (notably Axel Gandy). Again, I generally provided high-level coordination, though became more significantly involved in a subset of work.

e. Genetic analysis of SARS-CoV-2 evolution and variants: Genetic analysis was led by Erik Volz throughout, though with the rise of the Alpha variant, we shifted substantially more people to this work stream, such that I had a moderate level of day-to-day technical involvement from December 2020 onwards.

25. From late March 2020 onwards, these workstreams were coordinated via a management group consisting of myself, Azra Ghani, Samir Bhatt, Katharina Hauck, Natsuko Imai (our WHO liaison who also acted in a project management role), Sabine van Elsland (our communications manager) and other senior staff (*e.g.* Erik Volz, Nick Grassly) on a more ad-hoc basis. This met typically weekly throughout 2020, and then more infrequently thereafter. In addition, we held regular whole team meetings to allow ongoing research to be presented and critiqued; these occurred once or twice a week throughout much of 2020, then less frequently from 2021 onwards.

26. Staffing of COVID-related research varied substantially over the course of the pandemic: while only five people contributed to our initial analysis of the scale of the Wuhan epidemic [NF/003

- INQ000236275], this number scaled rapidly so that approximately 80 staff and students were contributing by mid-March 2020. The number then gradually declined thereafter, such that approximately 20 full-time-equivalent staff and students were working on COVID-19 research by the end of 2021.

27. A number of MRC GIDA staff and students were seconded to government departments (Cabinet Office, Government Office of Science and PHE/JBC/UKHSA) during the pandemic.

28. Some one-off research projects were also undertaken, led by individual academic staff – notably work on the role of testing and contact tracing, led by Nick Grassly [NF/004 - INQ000215948; NF/005 - INQ000262557]. MRC GIDA also contributed staff and technical expertise to other COVID-related research undertaken by Imperial College, including the REACT infection survey led by Paul Elliott, a number of epidemiological studies of household transmission led by Ajit Lalvani and the SARS-CoV-2 human challenge study led by Chris Chiu.

29. Appendix B provides a summary of all ICCRT outputs, grouped by type of output and topic. Individual outputs will be discussed below as appropriate.

D. Epidemiological analysis and modelling

Note: with the agreement of the Inquiry, this section was written with the assistance of Drs Charles Whittaker, Oliver Watson and Marc Baguelin

Introduction

30. The aims of epidemiological analysis and modelling of infectious diseases are to understand the spread of pathogens and associated health burden on populations, and to inform public health strategies for control and elimination.

31. The distinction between epidemiological *analysis* and *modelling* is a not precise. Analysis (e.g. estimation of disease severity, or of vaccine effectiveness) makes use of sophisticated statistical models (e.g. linear regression, survival models) which embed assumptions (such as the linearity of relationships between variables) often just as much as dynamical mathematical models of the type used to simulate epidemics. However, epidemic models attempt to represent more of the mechanistic process underlying disease transmission than purely statistical models, which focus on representing the correlations between variables.

32. All epidemiological models utilize data provided by surveillance systems and research studies to estimate epidemiological parameters, infer risk factors determining the spread or severity of the disease, and to give insight into the actual or potential effect of interventions. Outputs can support public health decision-makers, who use them to allocate resources, design and implement public health policies, and respond to outbreaks in real-time. Additionally, epidemiological analysis and modelling also provides insights into the long-term trends of infectious diseases, such as changes in disease burden and risk factors, and can help inform future prevention and control efforts.

33. The history of epidemiological modelling dates back to the 18th century, when the mathematician Daniel Bernoulli created a probabilistic model to assess the population benefit of the inoculation of smallpox. However, the history of modern epidemiological modelling is generally traced back to the early 20th century with the work of Ross (1911) and Kermack and McKendrick (1926-1927), who began studying the spread of transmissible diseases through populations over time. They developed mathematical models to better understand the dynamics of infectious diseases and the number and distribution of infections. Their work shared similarities with that of Lotka (one of the founders of theoretical ecology) and Volterra, who used mathematical models to study ecological systems such as predator-prey and host-parasite relationships. Since then, the discipline of epidemiological modelling has developed enormously, models have become hugely more sophisticated, but all current models still embed the core concepts included in those early models.

34. The very rapid growth in computing power seen over the last 40 years has driven the practical application of modelling to inform disease control, enabling models to incorporate much more detail (such as the structure of the human population). Perhaps even more importantly, that growth in computer power has allowed modern Bayesian statistical methods to be applied to epidemic models to allow rigorous estimation of epidemic model parameters by fitting models to epidemiological data.

35. Hence modelling has played an increasingly important role in informing public health responses to both major endemic pathogens (e.g. HIV, TB, malaria) and to emerging infectious disease threat (e.g. SARS-1 in 2003, H1N1 pandemic influenza in 2009, Ebola, Zika, MERS-coronavirus). More specifically, real-time modelling can enhance situational awareness and support control policy planning.

36. In emerging infectious disease epidemics, such as COVID-19, early epidemiological analysis tries to address four questions:

    a. *How far has it got?*

       What's the true scale of the epidemic? What proportion of cases are being missed?

    b. *How fast is it spreading?*

       What is the epidemic growth rate, how fast is international spread?

    c. *How bad is it?*

       How severe is the infection? What health burden might an epidemic cause?

    d. *What can we do?*

       What are the policy options, what impact and costs might they have?

Model types

37. Infectious disease models can be broadly grouped into three categories:

    a. *Statistical models*, which do not explicitly model the mechanism of transmission, but instead aim to estimate either summary statistics (e.g. epidemic doubling time) or epidemiological parameters (e.g. incubation period, infection fatality ratio (IFR), vaccine effectiveness, excess deaths, properties of new variants). Because such models don't represent the transmission process, they are not well-suited to predicting medium or long-term epidemic trajectory, disease burden or the impact of interventions on transmission.

    b. *Mechanistic transmission models* (often called epidemic models), which explicitly account for the underlying mechanisms of disease transmission and aim to identify the drivers of transmissibility. They make more assumptions about disease dynamics and require transmission parameters to be estimated or assumed. Such models can be used to estimate epidemiological parameters such as the reproduction number (R), but their most useful application is examining the potential impact of interventions, especially on transmission.

    c. *Semi-mechanistic models*, which are a blend of both statistical and mechanistic models – they represent some aspects of transmission dynamics (mathematically, by representing an epidemic via a renewal equation or branching process) but in a simpler way than full epidemic models. They are best suited to estimating epidemiological parameters such as R and making short-term epidemic projections.

38. Mechanistic transmission models can further be divided into two broad categories:

    a. *Individual-based models* (also called agent-based models or microsimulations), which explicitly represent every individual in a population, and can therefore be arbitrarily complex in how population structure (e.g. inclusion of households, schools and workplaces), disease transmission (e.g. inclusion of super-spreading) and progression (e.g. variation in time to or probability of hospitalisation) is represented. Individual-based models have the disadvantage of being relatively computationally costly, due to the computation required to track and evolve the states of millions of individuals.

    b. *Compartmental models*, which only track the number of individuals in the population in certain categories or states (e.g. susceptible, exposed, infectious and recovered – SEIR) and the flow of people between these states. While such models can stratify populations by characteristics such as age and/or region of residence, they represent the population (and therefore the transmission process) in a coarser-grained manner than individual-based models. However, compartmental models have the major advantage of being much less computationally complex to run, making them more suited to repeated real-time use in an epidemic. They also tend to have fewer parameters which need to be estimated or assumed than individual-based models.

39. While individual-based models always simulate an epidemic as a stochastic (*i.e.* random) process (meaning that no two simulations are precisely identical), compartmental models can be coded to be either deterministic (approximating epidemic spread as the average flow of people between states) or stochastic (representing the randomness of transmission and other processes). Stochastic models are more realistic, since transmission is a random process in reality, but are more computationally costly to run, especially given multiple model runs are typically needed to calculate average trends. Deterministic models provide a good approximation of disease dynamics when random fluctuations are small – *i.e.* when case numbers are high and a large population is being modelled.

40. Individual-based models allow for a more nuanced and detailed simulation of the spread of disease (including tracking disease spread from individual to individual), as well as more precise representation of individual-level factors (such as households and workplaces that an individual spends their time within). As such, they are better suited to modelling non

pharmaceutical interventions (NPIs) such as case isolation within the home, contact tracing or working from home.

41. Compartmental models do not capture the complexities and heterogeneity of individual behaviour, and so mostly have to represent all NPIs in the same way – as reductions in population contact rates or a change in R.

42. Epidemiological modelling is a far more multi-disciplinary exercise than might be expected; to develop and parameterise a transmission model, information is needed on:

    a. The natural history of infection – how infection progresses in a person, and how variable this is.

    b. Clinical severity and healthcare burden – the proportion of infected individuals who will be symptomatic, ill enough to require hospitalisation or ICU, the proportion who will go on to die – and information on the risk factors for severe outcomes (e.g. age).

    c. Transmission patterns – estimates of R from the epidemic itself, plus more detailed data on transmission derived from epidemiological studies (e.g. household studies, infection surveys).

    d. Demography – the age distribution of the population, population sizes locally, regionally or nationally (depending on the geographic scope of a model). For individual-based models, data on household composition, workplaces and schools are also needed.

    e. Population behaviour – most notably data on "who contacts who", typically derived from contact surveys. Data on population mobility (e.g. from mobile phones) can also be used as a proxy of overall contact rates. Behavioural surveys can provide information on the propensity to seek testing, wear masks or adhere to voluntary guidelines or mandatory measures.

    f. Intervention effectiveness – treatments, vaccines and NPIs. Randomised clinical trials provide the gold-standard estimates, but observational studies are equally important, especially for NPIs (given these have rarely been evaluated in randomised trials).

    g. Immunology – to characterise the dynamics (e.g. duration) of infection- and vaccine-induced immunity and the level of protection provided against infection, mild disease, severe disease or death.

h. Viral genetics and evolution – new variants can have different properties, which need to be captured in models.

43. During the COVID-19 pandemic, we developed all three categories of models, including both deterministic and stochastic compartmental models (see section E).

Retrospective versus prospective modelling

44. Infectious disease models can be applied either retrospectively or prospectively:

a. *Retrospective modelling* involves fitting models to past data to analyse epidemic patterns, calculate epidemiological parameters such as the reproduction number, and evaluate the effect of prior interventions (e.g., determining the impact of non-pharmaceutical interventions).

b. *Prospective modelling* involves projecting future epidemic trends based on current information, often drawing on parameters and trends estimated using retrospective modelling.

45. There are two major types of prospective modelling:

a. *Prediction/forecasting* is a structured form of prospective modelling that involves making statistically rigorous predictions based on a set of assumptions, with the goal of making accurate forecasts. Typically forecasts for infectious diseases such as COVID-19 are only made over a time horizon of a few weeks, since epidemic trajectories (akin to the weather) are non-linear phenomena which tend to be poorly predictable over longer horizons due to, for example, changes in population behaviour, government interventions or even climate. Models used for forecasting can be mechanistic, semi-mechanistic or statistical.

b. *Scenario modelling* involves examining multiple medium- to long-term scenarios concerning the future of an epidemic, often exploring the potential outcomes of a menu of policy alternatives requested by policymakers. For example, scenario modelling was used in the COVID-19 pandemic by SPI-M-O to project how healthcare demand might be affected by proposed government interventions [NF/006 - INQ000262592]. Mechanistic models are needed for scenario modelling.

46. The distinction between scenario modelling and forecasting is important and largely reflects the distinct types of questions that each approach is best suited for. Scenario modelling answers "What if" questions, such as "What would ICU admissions be if new NPIs were

introduced next month?". Forecasting focuses on "What will" questions, such as "What will daily COVID-19 cases be in a week?". Scenario modelling should not be viewed as formal prediction; in reality, actual government policies rarely if ever precisely match what was modelled, and precise prediction of the impacts of policy and population behaviour changes is generally not possible.

47. Research spanning many application areas of modelling (beyond just infectious diseases) has demonstrated that more reliable short-term predictions are obtained by using an ensemble of multiple different models than any single model can typically provide. Such approaches were adopted by a number of countries in the COVID-19 pandemic to provide short to medium term forecasts or projections of trends in cases, hospitalisations and/or deaths (e.g. in the US). This approach was used by SPI-M-O during the pandemic to generate R estimates and medium term (typically <4 week) projections. Ensemble approaches combine probabilistic estimates or projections made with multiple different models (for SPI-M-O, typically 8-12) in a statistically principled manner. SPI-M-O published an assessment of its approach to projections in December 2020 [NF/007 - INQ000191093]. Formal ensemble methods are less commonly used for scenario modelling, but more qualitative comparison of the results from multiple different models was standard practice for SPI-M-O throughout the pandemic when modelling intervention options or long-term scenarios.

Uncertainty, parameterisation, validation and verification

48. Uncertainty is inherent to epidemiological analysis and modelling. It is present across every aspect of the modelling process; from collection of the data used to calibrate the model, to design choices around model structure, to generation and interpretation of model outputs. Modellers must take into account numerous sources of uncertainty, including data limitations (*e.g.* case underreporting due to limited testing capacity), the uncertainty in epidemiological parameters arising from limited data (such as the basic reproduction number) and the inherent stochasticity in how pathogens spread through populations. These sources of uncertainty mean that modellers frequently make probabilistic projections which incorporate uncertainty in model parameters in a statistically principled manner. Where a full probabilistic analysis isn't feasible, sensitivity analysis can be used to show how model outputs vary depending on the values assumed for key parameters (or sometimes structural assumptions) for which there is significant uncertainty. It should be emphasised that both approaches are ways of *presenting* model uncertainty; neither *reduces* uncertainty. Formal forecasting is now

most commonly undertaken in a fully probabilistic framework (e.g. the SPI-M-O medium term projections), as this also allows results from different models to be more easily integrated into an ensemble forecast. Scenario modelling and other forms of illustrative modelling (e.g. to assess the potential effect of single policies such as bubbles) can make use of probabilistic methods (i.e. providing probability distributions for all model outputs, rather than single values), or use sensitivity analysis; both approaches were commonly used in SPI-M-O modelling throughout the pandemic.

49. As simplified representations of complex social and biological systems, epidemic models have intrinsic limitations. For instance, models typically represent human behaviour as constant over time, modified only by interventions. In reality, behaviour has changed over the pandemic in response to government policy, messaging and public perceptions of risk – often in unpredictable ways. Very few models attempt to predict such changes, largely because we don't have a good quantitative and predictive understanding of them.

50. Uncertainty is highest during the early stages of an epidemic, when data availability is most limited. This is particularly the case for novel pathogens such as SARS-CoV-2 where the absence of previous outbreaks in human populations precludes use of historical data. This leads to significant uncertainty in estimation of key parameters in epidemiological models, such as the basic reproductive number, the incubation period, and disease severity. However, as more data become available, epidemiological models can be refined and uncertainty reduced, allowing for more precise modelling and informed decision-making.

51. During an epidemic, major changes in interventions, population behaviour or the pathogen can introduce additional uncertainty. Examples include the implementation of new (previously untested) NPIs, external events that affect behaviour (such as the 2020 Euros football tournament leading to a surge in cases), new technologies (such as lateral flow testing or the Covid app), new therapeutics (such as vaccines), or changes in the biology of the pathogen (such as the emergence of a variant with distinct epidemiological characteristics).

52. Uncertainty in epidemiological modelling is typically larger for prospective modelling (looking forwards in time) than retrospective modelling (looking backwards). Prospective modelling necessitates making assumptions about future events or trends which are inherently uncertain. It is also necessarily based on data from the past and therefore relies on an assumption that past dynamics are representative of expected future dynamics – an

assumption that can be invalidated by new variants or major changes in policy or population behaviour.

53. Modellers try to reduce uncertainty by obtaining the most precise estimates of model parameters possible, using a wide range of data sources. These include behavioural data (describing population-level patterns of behaviour relevant to pathogen transmission such as the degree of adherence to control measures); demographic data (important for diseases like COVID-19 where severity varies with age); and data from epidemiological studies (e.g. on the incubation period). Where possible, parameters are estimated from such independent sources, but where this isn't possible, unknown parameters can be estimated by fitting epidemic models to surveillance data - this "tunes" parameters in a statistically principled way (e.g. using Bayesian methods) to find values which allow a model to best reproduce epidemic trends.

54. A particular challenge for epidemic modelling is that epidemics are only partially observed; recorded case counts only represent a fraction of total infections, since not everyone develops or recognises symptoms, or seeks testing. Data on infection *prevalence* is therefore particularly valuable – the proportion of the population who have been infected thus far in an epidemic (measured through serological surveys) or the proportion who are infected now, irrespective of symptoms (measured through studies such as the ONS or REACT infection surveys). Such data allow estimation of the degree to which recorded cases underascertain infections, allow tracking of the accumulation of immunity in the population, and allow reliable estimates of the IFR and IHR to be made.

55. Model verification is testing whether the code used to represent a model is formally correct (*i.e.* gives the correct solution to the equations used to specify the model). In reality, as is the case for most mathematical modelling in science, formal verification (in the computer science sense of the term) is rarely feasible for all but the simplest epidemic models. Therefore, a number of more heuristic approaches are used to check model implementations. One is to compare the outputs of different models when each is configured with the same parameters. A second is to examine model output for simplified "edge cases" where model output can be compared with analytical calculations or solutions of much simpler models. Another – for models which are designed to be fitted (using Bayesian methods) to data – is to test the that model fitting code produces unbiased estimates using simulated data.

56. Moreover, the last 20 years have seen increasing professionalisation of research software development, assisted by version control, unit testing and the advent of tools such as GitHub – and coinciding with a trend towards teams rather than individuals developing code, open source development and code modularisation and reuse. ICCRT has a professional research software engineering (RSE) team of some 10 staff, many of whom contributed substantially to the development of the new COVID-19 models (see section E) we created in response to the pandemic.

57. If time and resources permit, model code can also be reviewed by independent groups. However, this is uncommon, and rarely even forms part of peer-review of scientific papers making use of model (statistical or mechanistic) output. That said, the pandemic has seen an acceleration of pre-existing trends for journals to require that code and data are published with scientific papers to allow results to be reproduced. Except where prevented by data protection issues (associated with the analysis of individual-level data), ICCRT have followed this principle with our peer-reviewed journal articles on COVID-19 and the majority of our reports and preprints. All our COVID-19 epidemic models are open source.

58. Model validation is a less precisely defined term than verification, but is generally the process of checking that a model provides an "adequate" description of the system the model is designed to represent. Often this involves fitting a model to surveillance data (mentioned above). Model fitting is a statistical procedure involving exploring a large range of parameters to determine the sets of parameters that are most consistent with the observed data. The plausibility of the resulting parameter estimates and the model's ability to accurately capture the epidemiological patterns present in the data can then be assessed. Other methods of validation may include external validation, which involves the comparison of model outputs to independent data sources not involved in model fitting – so-called "out-of-sample" validation. In general, true out-of-sample predictive validation of epidemic models is challenging, since no two epidemics are identical (e.g. COVID-19 pandemic trajectories in different European countries), meaning some parameters always need to be re-estimated. However, for prospective analyses focussed on forecasting, assessment of past predictions from previous rounds of forecasting (but not scenario analysis) can be used for model selection and to refine models.

59. When modelling is being used to inform policy-making and model development is occurring under severe time pressure, an additional check on the robustness of model outputs or policy-

relevant conclusions is to compare the outputs of independently developed models. This is the principal approach to model validation adopted by SPI-M-O. Throughout the pandemic, comparable modelling was requested from multiple academic groups; never less than two, and nearly always 3 or more.

Uses of statistical and semi-mechanistic models

60. Early in an emerging infectious disease epidemic, statistical models are used to estimate epidemiological parameters such as the incubation period, serial or generation interval, symptomatic fraction, the infection hospitalisation ratio (IHR), the case fatality ration (VFR), and the infection fatality ratio (IFR) – and to understand how those parameters vary over time and with risk factors such as age or pre-existing medical conditions.

61. Statistical or semi-mechanistic models are used to estimate epidemic growth rate (*i.e.* doubling time) and/or R.

62. In general, the choice of model utilised depends on a combination of the exact research question being asked and the data available.

63. An important distinction is whether available data is at the individual-level (i.e. data on every case) or aggregate (e.g. counts of cases per day). Individual-level data are essential for estimation of the incubation period, serial interval and other delay distributions. They are also important for evaluating differences between groups of cases – such as assessing how disease severity varies with age, sex, ethnicity, vaccination history or the variant of virus causing the infection.

64. Individual-level data is also essential for analyses which rely on data linkage – for instance, evaluation of how disease severity (IHR and IFR) varies by virus variant, or estimation of vaccination effectiveness (VE). In the COVID-19 epidemic, many UKHSA and SPI-M-O analyses relied on (anonymous) linkage of multiple large databases of COVID-19 testing records, hospital episode records, vaccination records, deaths and viral genetic sequences.

Mechanistic epidemic models for respiratory viruses

65. Mechanistic transmission models for respiratory viruses such as COVID-19, influenza, respiratory syncytial virus (RSV) and measles typically have the same basic structure. These models aim to describe the spread of the disease in a population by tracking the infection and

disease states of either groups of individuals in a population (in the case of compartmental models) or individuals (in the case of individual-based models).

66. In all respiratory virus models, transmission of the pathogen is driven by a combination of the rate at which individuals make contact with one another, the probability of successful transmission upon contact and duration for which individuals remain infectious.

67. Because the transmission route is the same for all respiratory viruses, the same population contact rate data (on social contacts) is used to parameterise influenza, measles and COVID-19 models. Transmission models do not represent the physics of transmission, and hence do not distinguish between, for instance, droplet versus aerosol transmission. However, models of diseases with very different transmission routes (e.g. sexually-transmitted or water-borne) do differ in structure from respiratory disease models.

68. Hence the fundamental approach to modelling different respiratory viruses is largely the same, with the differences between models of different pathogens primarily being in how a model is parameterized – with R_0, the incubation period, infectious period, infection hospitalisation ratio (IHR) and infection fatality ratio (IFR) being key parameters, plus how some of these parameters (e.g. susceptibility, symptomatic proportion, IHR and IFR) vary with age and by other population characteristics. Hence a model coded to model COVID-19 can be repurposed to model the next influenza pandemic, so long as parameter values are appropriately updated. The converse is also true.

69. Respiratory virus models are parameterised to represent contact patterns (and sometimes networks) relevant for respiratory virus transmission. Contact patterns refer to the frequency and duration of interactions between individuals, while networks describe the relationships between individuals in a population. Understanding these is essential for predicting the potential spread of a virus and for informing effective control measures. Models that take into account the specific ways in which individuals interact, can potentially (if they can be parametrised) provide more accurate predictions of transmission patterns and the effects of NPIs than those which more crudely capture overall contact rates.

70. When designing models, choices are often required about the level of granularity to include. These decisions are often based on the specific questions a modeller is aiming to answer and the availability of data. For example, if the goal is to forecast hospital demand, a model may need to include explicit information about the passage of COVID-19 patients admitted to

hospitals and the time spent in different types of facilities (*e.g.* general ward versus ICU). The availability of data to parameterize models also plays a role in determining the level of granularity, with the absence of data often limiting the complexity of models that can be reasonably justified. For example, extending models to differentiate droplet versus aerosol transmission (or transmission associated with indoor versus outdoor contacts) would introduce additional model parameters that currently are unable to be reliably estimated from available data. In such cases, it is often better to use simpler models with more coarse-grained representations of contact processes for which data are available.

71. Population heterogeneity (*i.e.* differences between people) is the type of granularity most frequently considered by modellers. Appropriately representing heterogeneity is important for accurately representing transmission and disease risk. Models can represent person-to-person variation in characteristics such as contact rates, infectiousness, susceptibility, and severity. Nearly all respiratory virus epidemic models allow model parameters to vary by age, given age is a predictor of both contact rates and clinical outcome. Compartmental models are limited in their ability to represent heterogeneity (though regional variation is often included), while individual-based models have fewer constraints. A type of heterogeneity which was important early in the pandemic was variation in the number of secondary infections an infected person generated (so-called super-spreading), and a range of individual-based models were developed to examine the implications of this for control (*e.g.* [NF/008 - INQ000191098; NF/009 - INQ000212222]).

72. In a long-running epidemic, it is common to extend and update models to incorporate new evidence as it becomes available. For example, while we were aware that reinfection occurs with endemic (and antigenically diverse) human coronaviruses typically every two years, the expectation that infection would generate good protective immunity for at least a year meant that early models did not include the waning of immunity. As data became available on the duration of immunity induced by COVID-19 infection in late 2020, models were updated to account for this. Additionally, models may need to be extended to be able to model the effect of new interventions, such as vaccines and novel therapeutics. Furthermore, the emergence of new variants of a virus with distinct epidemiological properties requires models to be extended to model more than one variant at once circulating in the population, each with sometimes subtly different epidemiological parameters. All such extensions increase model complexity and the computational cost associated with running models.

73. Models for endemic diseases (such as seasonal influenza or measles) differ from those used to model the initial epidemic of a new virus. At the start of the pandemic, there was no immunity in the human population, meaning spread occurred between individuals of all ages. At the opposite extreme, measles only infects young children because everyone else has (lifelong) immunity. Thus, to model measles, it is essential for models to include births (and deaths), given it is newly born children who renew the susceptible population the virus can then infect. This type of dynamics has not yet been important for COVID-19, so few COVID-19 models include new births into the population (or other demographic processes such as migration). I would also comment that endemicity for COVID-19 will be more complex than for measles, given the rate of viral evolution and that immunity wanes over time.

The reproduction number, R

74. The basic reproduction number, R_0, is a metric that quantifies the average number of new cases generated by a single infected individual in an uninfected (entirely susceptible) population (i.e. at the start of an epidemic). It is determined by the combination of the rate at which individuals contact one another; the length of time individuals remain infectious, and the probability of pathogen transmission per contact event. It determines the potential for an epidemic to occur, as a spread will only continue and grow in a self-sustaining way if R_0 is greater than 1 and will diminish if R_0 is less than 1.

75. As immunity builds up in a population (due to infection or vaccination) and/or control measures (e.g. social distancing) are implemented, transmission intensity is reduced and we quantify this by the so-called time-varying reproduction number, R (also labelled R_t or R(t), where t is time) – the average number of new infections generated by a single infected individual at that stage of the epidemic. If R is above 1, the epidemic will be growing, while if it is below 1, daily case numbers will be declining. By tracking the changes in R, public health officials can monitor the effectiveness of control measures and make informed decisions on how to respond to potential surges in cases.

76. If the aim of policy is to suppress transmission and hence cause daily case numbers to decline, interventions need to reduce R to below 1. This means that R_0 also determines the control effort required for suppression. For example, if R_0 is 2, 50% of transmission needs to be blocked to get R to 1, while if R_0 is 4, a 75% reduction in transmission is required.

77. R solely quantifies epidemic growth rate; it says nothing about incidence measures such as the number of daily new cases or hospital admissions.

78. R can be estimated in a variety of ways, but most commonly using methods that analyse trends over time in disease indicators – such as incidence (reported confirmed cases, hospitalisations, or deaths) or prevalence (as measured by the ONS and REACT infection surveys). Estimation is complicated by several challenges during an epidemic. These include reporting delays (which make recent case counts appear lower than the true value and lead to erroneous conclusions of epidemic decline unless corrected for), and changes in testing capacity and propensity to test over time (which alters the fraction of infections being captured by surveillance systems over time and which can therefore present erroneous pictures of epidemic growth). Furthermore, in the case of an emerging infection such as SARS-CoV-2, the natural history of the pathogen is often not well understood at the start of the epidemic, which adds to the uncertainty in estimates of R. This is because estimating R requires good estimates of a parameter called the *generation interval*, which measures the distribution of time between when cases become infected and when they transmit to others.

79. R can only be estimated retrospectively; COVID-19 cases recorded today were likely infected a week or more ago, and new admissions to hospital were likely infected even longer ago. Hence estimates of R based on current data reflect transmission in the past. Given the range of data sources used by SPI-M-O groups to estimate R during the pandemic, and the 3-4 days it took for new R estimates to be signed-off, most of the weekly official R estimates released by the UK government actually quantified transmission trends 2-3 weeks earlier.

E. Imperial COVID-19 models

Note: with the agreement of the Inquiry, this section was written with the assistance of Prof Samir Bhatt and Drs Anne Cori, Daniel Laydon, Oliver Watson and Lilith Whittles

Introduction

80. There was never a single "Imperial Model". Over the course of the pandemic, ICCRT staff developed a variety of mechanistic and semi-mechanistic models of COVID-19 transmission. The structure and focus of each model differed depending on its purpose, or the input data sources to be used for model calibration. Despite these differences, model parameters (e.g.

generation interval, disease severity) were kept consistent across all models to the extent possible given the differences in model structures.

81. In addition to the mechanistic and semi-mechanistic models discussed in this section, ICCRT developed many statistical models for specific epidemiological analyses during the pandemic.

82. Model parameters (notably R_0, the generation interval, incubation period distribution, IHR and IFR by age, hospital stay by age, symptomatic proportion by age, how susceptibility varied by age, waning of immunity) and structures were regularly updated as more data accumulated, in response to changes in the epidemic (e.g. new variants), as new data streams became available, and as scientific understanding of COVID-19 improved. . In addition, models were extended (e.g. to include vaccination) to allow them to address new commissions from policymakers and public health partners.

83. Model validation and verification was a priority throughout the pandemic, and we adopted a variety of approaches to this. First, models were fitted (typically using Bayesian methods) to empirical data wherever possible. Second, models were used as sense-checks for each other: consistency of outputs, given differing model structures and code, adds confidence that each is giving intended results. Our model outputs were also compared with those of other research groups, with differences between groups being discussed (e.g. in SPI-M-O) to ensure they arose from genuine differences in modelling choices, assumptions or parameter estimates and thus were a valid reflection of modelling uncertainty.

84. Changes to model code were reviewed by other team members before acceptance and also checked using "regression tests", which ensure the addition or refinement of model features preserves previous output and functionality. As is standard practice, we conducted such tests at every stage of model development for every model. Often, models are implicitly nested, such that it is possible to turn off features of one model so that it closely resembles another. For example, turning off stochasticity in SirCovid, or removing the households, schools and workplaces in CovidSim, should yield results highly similar to a simpler structured SEIR model, for which model behaviour is better understood and more predictable. Of course, this approach mutes some functionality and purpose of the more complex model, but is nevertheless important to ensure models behave as they should. Finally, all of our model code is open-source, allowing outside scrutiny.

85. The table on the following page summarises each of the semi-mechanistic and mechanistic transmission models used by the ICCRT team, after which I provide a description of each model in turn.

Table: Summary of ICCRT COVID-19 transmission models

	EpiEstim	Epidemia	SirCovid	squire/nimue	Daedalus	CovidSim	safir
Type	Semi-mechanistic	Semi-mechanistic	Mechanistic-compartmental	Mechanistic-compartmental	Mechanistic-compartmental	Mechanistic-individual based	Mechanistic-individual based
Stochastic or deterministic	Probabilistic	Probabilistic	Either	Either	Deterministic	Stochastic	Stochastic
Computational requirements	Low	Moderate	Moderate	Moderate	Moderate	High	High
Spatially-explicit	×	×	×	×	×	✓	×
Fitted to data / used for inference	✓	✓	✓	✓	×	After Report 9	Uses squire fits
Used for scenario modelling	×	✓	✓	✓	✓	✓	✓
Age-specific inputs & outputs	×	×	✓	✓	✓	✓	✓
Age-specific contact patterns	×	×	✓	✓	✓	✓	✓
Modelling of multiple severity levels (e.g. infections, cases, hospitalisations, deaths)	×	✓	✓	✓	✓	✓	✓
Asymptomatic/presymptomatic transmission	×	×	✓	✓	×	✓	✓
Includes care home sector	×	×	✓	×	×	After Report 9	×
Models households, schools and workplaces	×	×	×	×	×	✓	×
Can model multiple variants	×	×	✓	✓	×	×	✓
Includes vaccination	×	×	✓	✓	✓	✓	✓
Includes waning immunity and reinfection	×	×	✓	✓	×	×	✓
Detailed immunological model	×	×	×	×	×	×	✓
Includes economic impacts	×	×	×	×	✓	×	×
NPIs included	Time varying R (fitted)	Time varying R (fitted), inference of individual NPI effectiveness from timing, pooling across regions	Time varying contact rates (fitted)	Time varying contact rates (fitted)	Time varying contact rates by economic sector (optimised)	Case isolation, household quarantine, contact tracing, social distancing, shielding, school/ workplace closure	Time varying contact rates (fitted)

Shared representation of healthcare demand and disease severity

86. The mechanistic models listed in the table above (SirCovid, squire, Daedalus, CovidSim and safir) all share a very similar representation of COVID-19 disease progression and associated healthcare demand, represented graphically in the figure below:

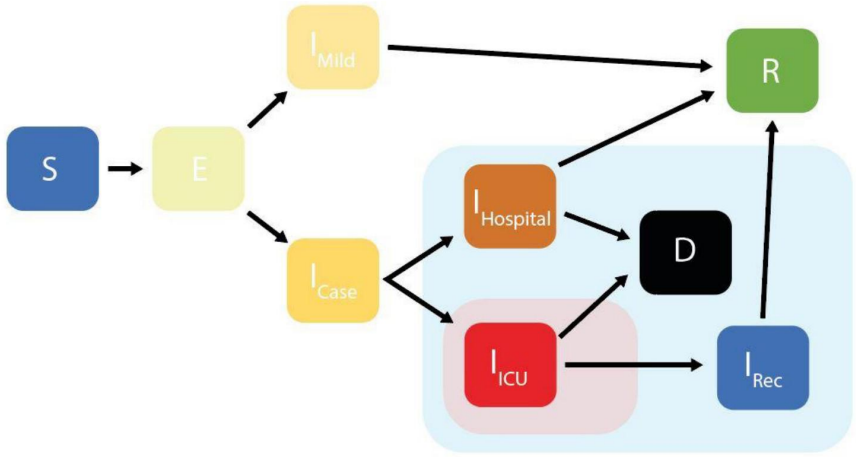

Figure: Model structure shared by mechanistic COVID-19 transmission models. S = Susceptibles, E = Exposed (Latent Infection), I_{Mild} = Mild Infections (all infections not requiring hospitalisation), I_{Case} = Infections Requiring Hospitalisation, $I_{Hospital}$ = Hospitalised, I_{ICU} = ICU, I_{Rec} = Recovering from ICU Stay (Requires Hospital Bed), R = Recovered, D = Dead. Pale blue box indicates hospitalisation, pale red box indicates ICU.

EpiEstim

87. EpiEstim is a generic semi-mechanistic model (based on renewal equations) MRC GIDA developed before the pandemic to estimate how R changes over time in an epidemic. It is not specific to any pathogen, and requires only two inputs: 1) an incidence time series (e.g. the number of new cases per day) and 2) the probability distribution of the serial interval (the time between symptom onset in a case and symptom onset in their infector). EpiEstim then translates observed epidemic growth over calendar time into estimated growth per generation, as measured by the effective reproduction number R_t, defined as the number of secondary cases per index case at time t. The advantages of EpiEstim are:

a. it is not specific to a pathogen or an epidemic context

b. it relies on only two inputs, which are both typically available through traditional epidemic surveillance

c. it is available in ready-to-use software

d. it is computationally efficient.

88. The model has limitations. First, it interprets the growth in observed cases as the true epidemic growth - inconsistent reporting rates may bias estimates. Second, EpiEstim assumes no heterogeneity in contact rates (e.g. between age groups), and so cannot directly elucidate transmission drivers. Third, although EpiEstim enables detection of changes in transmissibility (i.e. R) over time, it offers no insight into the cause of these changes.

89. Its simplicity and low data requirements mean that EpiEstim is a useful first model to use early in an epidemic, before detailed information is available. Many researchers and public health institutions globally have used EpiEstim to monitor COVID-19 transmissibility in real-time. In the UK, some (non-ICCRT) models submitted to SPI-M-O used EpiEstim [NF/010 - INQ000223965]. In our COVID-19 work, we have used it:

a. as a benchmark against which to validate more complex models [NF/011 - INQ000212077].

b. to characterise the extent to which global mobility data could predict SARS-CoV-2 transmission levels [NF/012 - INQ000262559; NF/013 - INQ000262606],

c. to provide real-time weekly global forecasts of COVID-19 deaths [NF/014 - INQ000262588; NF/015 - INQ000273322],

d. to estimate the effective transmission advantage of emerging variants [NF/016 - INQ000262562; NF/017 - INQ000262591]

90. Prompted by the emergence of the Alpha variant in early 2021, we extended EpiEstim to estimate the transmission advantage of new variants [NF/016 - INQ000262562], and applied this extension to estimate of the transmission advantage of the Delta variant [NF/017 - INQ000262591].

Epidemia

91. Epidemia has similar aims to EpiEstim – estimating how R changes over time from time-series data – but differs from EpiEstim in modelling the (latent or unknown) underlying infection process, and not just the occurrence of reported cases. It therefore accounts for infections that may not have been included in reported case data. Epidemia also adjusts for lags in the

data such as the time between infection, onset of symptoms or possible death. An advantage of epidemia is its rigorous statistical treatment of uncertainty, at a cost of being more computationally demanding than EpiEstim. Epidemia's data inputs are time series of one or more of: cases, hospitalisations, and deaths, as well as the timing of interventions, such as school closure dates. Epidemia also requires epidemiological parameters such as the generation interval and IFR. Like EpiEstim, Epidemia's mathematical foundation is in renewal processes.

92. Epidemia can also be used to retrospectively estimate the effects of interventions on transmission through a parametric form of R. This uses a generalised linear model for R with fixed and random effects. Fixed effects model known factors, such as the timing of NPIs. Random effects are a statistical technique to account for unknown factors, such as behavioural changes.

93. Epidemia was extended as the pandemic progressed. Its first application was to estimate the effect of NPIs (among other results) across 11 European countries [NF/018 - INQ000262556; NF/019 - INQ000228162], work which was later considerably extended by other groups (e.g. Brauner et al 2020 [NF/020 - INQ000262594]). It was then applied to track state-level transmission and NPI impact in the USA [NF/021 - INQ000262558], and to compare the responses of the UK, Denmark and Sweden during the first wave of the pandemic [NF/022 - INQ000262573]. It was used in a web platform we developed to track the progress of the UK pandemic by local authority area, giving real-time estimates of R and short term projections [NF/023 - INQ000262564]. Last, it was applied to an examination of the effects of NPIs in the second wave of the pandemic [NF/024 - INQ000218211].

SirCovid

94. While models such as EpiEstim and Epidemia are useful for generating R estimates, since they don't model the full transmission process (*i.e.* infectious people contacting susceptible individuals), they are less suited to projecting the potential impact of future intervention options. In addition, EpiEstim can only be fitted to a single datastream at once. Conversely, the computational requirements of largescale individual-based models such as CovidSim (see below) make them unsuited to be repeatedly fitted to (or calibrated against) surveillance data in real-time during an unfolding epidemic.

95. To address these limitations, in February 2020 we began development of an intermediate complexity model, SirCovid, for UK real-time modelling throughout the pandemic.

96. SirCovid is a stochastic (though it can also be run deterministically), compartmental SEIR (Susceptible, Exposed, Infectious, Recovered) model of COVID-19 transmission. It was designed to model transmission in each of the seven England NHS regions and three devolved administrations, to estimate R (for each region and overall) and to produce short- to medium-term projections. The model divides each regional population by age and infection status, and represents hospital clinical pathways (general, ICU and post-ICU bed occupancy) in some detail. NPIs were modelled as changes in the average contact rate, resulting in temporal changes in transmission. The model has been greatly extended as the pandemic progressed in response to new priorities. Modelling the care-home sector was added from May 2020, vaccination from October 2020, waning immunity from November 2020, and the modelling of multiple variants from January 2021.

97. SirCovid was designed from the outset to be extendable to include more data streams in the model calibration process as they became available. By the end of 2020, the model was being calibrated weekly using Bayesian methods against data on daily recorded deaths, PCR-confirmed Pillar 2 cases, hospital admissions, hospital bed occupancy, individual patient outcomes, contact surveys, infection surveys (such as ONS and REACT-1), and serological surveys.

98. The model is implemented as a set of packages in the R programming language, some of which make use of components written in C++ for performance. All packages developed for running and fitting the model are publicly available on the GitHub repository. We ensured reproducibility of all SPI-M-O modelling results by using another R package, *orderly*, to store every model run with the code, data and model parameters used to generate it. *Orderly* was written by the MRC GIDA research software engineering team before the pandemic to improve reproducibility in epidemiological modelling.

squire & nimue

99. squire is an age structured SEIR model of COVID-19 transmission which was developed to support COVID-19 modelling across a large number of low and middle income countries (LMICs), allowing easy specification of country-specific parameters such as contact matrices,

demography and health constraints [NF/025 - INQ000262629]. The model was based on SirCovid, using a similar structure and fitting process, but only uses mortality data (reported COVID-19 deaths or excess deaths).

100. The model can run deterministically or stochastically. The latter is used for running model projections, while the former is used to improve performance of model calibration/fitting to country-specific mortality data. This was necessary given that the model was initially used for providing daily epidemic fits to all countries in the world. Daily model fits and projections were made initially just for LMICs, before being extended to include high income countries later in 2020. These were produced as part of a project part of a project supported by the Wellcome Trust that was developed in close consultation with the UK Foreign, Commonwealth and Development Office (FCDO). Projections focussed on the healthcare demands each country may expect based on current epidemic trajectories. Additional scenarios were presented that provided projections under alternative scenarios in which transmission increased/decreased in response to change in NPIs.

101. We developed a website to present this modelling with the specific model outputs freely available for download. These outputs were used as inputs to a number of other projects, such as the WHO Essential Supplies Forecasting Tool, as well as the WHO Health Systems Governance and Financing division to mobilise resources for LMIC countries [NF/026 - INQ000262605]. To further assist country scenario planning, the output model fits were used within the covidsim.org web-based tool (different from the CovidSim model that was used in Report 9), which allowed users to explore alternative epidemic trajectories.

102. nimue was then developed as an extension to squire to model vaccination, allowing prioritisation of vaccination by age, risk-group and healthcare worker status to be specified. Multiple vaccines are considered by aggregating properties of each vaccine rather than explicit representation. Waning of vaccine efficacy and infection-induced immunity was incorporated, with later versions including provision for delivery of booster doses.. The model was adapted for variants of concern using a time-varying change to the hospitalisation rate, transmissibility, vaccine effectiveness and evasion of prior infection-induced immunity.

103. nimue supported WHO-EURO vaccine prioritisation recommendations and was used for the global assessment of the impact of COVID-19 vaccination. For this purpose, it was

extended to consider national excess mortality and therefore obtain more accurate estimates, given substantial under-ascertainment of cases and deaths in countries without adequate mortality reporting.

104. Reporting biases were also a significant concern when modelling the global spread of COVID-19. The data available for modelling was often limited, particularly in low- and middle-income countries, and the biases in reporting of COVID-19 deaths and cases were difficult to quantify. To explore these, several individual studies conducted in select locations were used to quantify the proportion of COVID-19 deaths that were undetected and were provided as oral evidence to the UK International Development Committee [NF/027 - INQ000262571]. As our understanding of under-reporting of COVID-19 mortality improved, our global modelling efforts focussed on fitting models to excess mortality data, which has been shown to better represent the impact of COVID-19 in LMICs.

Daedalus

105. Daedalus is a deterministic compartmental SEIR model of COVID-19 transmission coded in Matlab.. The model is very similar to both squire and SirCovid in design, except that it stratifies the modelled population (and contact rate matrices) into 63 economic sectors and incorporates a representation of different NPIs which not only describes the impact on disease transmission, but also on economic output. It allows NPIs to be targeted on specific sectors and includes optimisation code to allow intervention strategies to be optimised to maximise economic output subject to healthcare system demand constraints (e.g. keeping ICU occupancy below a certain threshold). The model has been applied in both high [NF/028 - INQ000212184; NF/029 - INQ000220361] and middle-income settings, and in assisting WHO planning.

CovidSim

106. CovidSim is a large scale (populations of up to hundreds of millions) individual-based transmission model of directly transmitted pathogens written in C/C++. It is adapted from earlier code used to model pandemic influenza which was then extended to model Ebola transmission in West Africa. Both model code and parameters were changed to reproduce

the epidemiology of COVID-19 – most notably in representing its impact on healthcare demand. The model of disease severity and healthcare demand embedded within CovidSim is the same as that used in SirCovid, though the added flexibility given by individual-based models means that more realistic representations of delay distributions (*e.g.* incubation period, symptoms to hospitalisation) are used than is possible in compartmental models. In addition, CovidSim represents person-to-person variability in infectiousness (*i.e.* super-spreading behaviour) – something which compartmental models are unable to do.

107. Unlike the models described above, CovidSim is spatially explicit, essentially taking a map of a given geography (e.g. Great Britain), dividing it into 1km by 1km cells, and populating those cells according to detailed demographic data. Each person is allocated to a household and either a school, workplace or university. Census data on household and workplaces size distributions, together with the age-distribution of the population, are thus used to generate a synthetic population mirroring the geography under discussion (GB and the US were the examples used in Report 9). Population mobility (i.e. travel from place to place) is represented using a so-called gravity model of fitted to commuting data.

108. Early in the pandemic, this model was used extensively to estimate healthcare demand and the potential impact of NPIs, most prominently in Report 9 [NF/001 - INQ000049647]. Because CovidSim explicitly represents social structures such as households, schools and workplaces (and the broader social network created by the overlap of those), it can be used to model the effects of NPIs targeting those structures in a more explicitly mechanistic manner than compartmental models (such as SirCovid) which model NPIs as changes to an overall contact rate. CovidSim explicitly represents the effects of case isolation, household quarantine, contact tracing (including app based contact tracing), social distancing, shielding and school or workplace closure (blanket or reactive to cases).

109. As with our other COVID-19 models, CovidSim was further developed and extended as the pandemic progressed. After Report 9 was released, we received support from Microsoft and github to refactor what previously had been relatively monolithic C code. The refactored code was open sourced (with regression tests and example parameter sets) on github in April 2020, followed shortly thereafter by parameter sets and scripts which allowed the results of Report 9 to be reproduced.

110. External groups verified reproducibility of the Report 9 results [NF/030 - INQ000262570] and applied CovidSim to further COVID-19 modelling [NF/031 - INQ000262581].

111. The Uppsala University analysis of intervention policy options in Sweden [NF/032 - INQ000262596] did not use the CovidSim model code or parameter files. Rather, that group developed their own individual based simulation model. Their model is no more "the Imperial model" than is any individual-based simulation of COVID-19 developed during the pandemic (of which there are at least 6, to my knowledge). Most individual-based models share features in terms of how they represent populations (*e.g.* including households, schools and workplaces), but can differ substantially in how they represent the epidemiology and transmission of COVID-19 and the impact of government interventions (voluntary or mandatory). To give another analogy, the Uppsala model is no more the Imperial model than the LSHTM and Warwick compartmental transmission models are the same as SirCovid – all of which have some similarities in structure and parameterisation, but also differ in many details.

112. In March-April 2020, digital contact tracing was incorporated in CovidSim to allow assessment of the potential effectiveness of isolation of the contacts of symptomatic people (not infectious people only), identified via a smartphone app. Further, in April 2020 we allowed NPIs to vary in their restrictiveness over time. Taken together, these amendments allowed us to explore a wide range of exit strategies from lockdown [NF/033 - INQ000280728]. Care homes were additionally added to the model in May 2020, given their prominence in the first wave of the pandemic in the UK.

113. While CovidSim is sufficiently computationally intensive to make model fitting challenging, simple "Latin hypercube" parameter sampling was used to implement likelihood-based model fitting in March 2020 and a more sophisticated parallel MCMC based inferential platform was developed in July 2020.

safir

114. To better capture the interaction between infection-induced and vaccine-induced immunity, we developed *safir*, another individual based model of SARS-CoV-2 transmission, written in the programming language R. The initial focus for the development of *safir* was to model the value of vaccine booster doses to mitigate the global impact of COVID-19. Given

the complex aspects of immune dynamics associated with vaccination and prior exposure, the flexibility that an IBM afforded over the compartmental *nimue* model was necessary to capture these dynamics more accurately. This flexibility provided to be necessary with the emergence of the Omicron variant, with the model being used since late 2022 to address questions related to immune escape and the role of hybrid immunity. The structure of *safir* is not spatially explicit (unlike CovidSim) and instead mirrors *nimue*, allowing *nimue*'s model fits to be used as inputs to *safir*.

115. *safir* incorporates a well-respected and highly cited mathematical model of the immune response to vaccination published by an Australian team. We applied this model to vaccine data from England to develop an "immune profile" that reflects the overall immune response. The model represents immunity levels decreasing over time with a bi-phasic decay profile and describes the relationship between immunity and protection against mild disease (similar to infection), hospitalization, and death using logistic relationships. The same approach is used to model immunity from infections, with an additive model considering the combined impact of both vaccines and infections at the individual level. This approach also allows for potential immune escape from new variants by reducing the immune levels to new variants.

F. Early epidemiological characterisation of SARS-CoV-2 (January-March 2020)

Wuhan outbreak size and international spread

116. While I had noted reports of a cluster of pneumonia cases of unknown aetiology in Wuhan a few days earlier, I first became concerned about the new outbreak in the week of 6th January 2020 – in particular with the release of genetic sequences of a novel coronavirus on 10th January. I attended a NERVTAG Skype meeting about the Wuhan virus on 13th January [NF/034 - INQ000023107]. This was the same day that Thailand reported the first case detected outside China [NF/035 - INQ000262597], an event which motivated me to start epidemiological research on the novel coronavirus within MRC-GIDA.

117. Our first priority was to obtain an estimate of the potential true scale of the Wuhan outbreak, given that it was statistically unlikely that a case would have been detected outside China without many more infections having occurred in Wuhan than the 41 cases then

reported. At the time, there was much uncertainty about the extent to which cases in Wuhan were infected by zoonotic exposure (via animal markets) versus human-to-human transmission. If there were hundreds to thousands of human cases, self-sustaining human-to-human transmission was much more likely to be occurring.

118. Our estimates of the outbreak size in Wuhan did not rely on epidemic modelling; rather we estimated the probability that an infected person in Wuhan would have taken an international flight using data on international passenger numbers and the population size of Wuhan. The cumulative number of cases detected outside China by a specific date divided by that probability gave an estimate of the outbreak size in Wuhan by the same date.

119. In the interest of timeliness, we published these simple analyses as reports on the MRC GIDA website on 17th (Report 1[3]) and 22nd of January (Report 2 [NF/036 - INQ000212192]). Report 1 concluded from the 3 international cases detected by that time, that there were approximately 1,700 human cases (uncertainty range: 400-4,500) in Wuhan City with symptom onset by 12th January (the last reported onset date of any case at the time the report was written). Report 2, based on 7 international cases, concluded that there were 4,000 (uncertainty range 1,000-9,700) cases in Wuhan with symptom onset by January 18th.

120. While these estimates compare well with the retrospective case count published by China CDC (on February 21st 2020) of 6,174 symptomatic cases with disease onset by 20th January [NF/037 - INQ000215642], they likely represented a minimum bound on the true number of infections in Wuhan in January 2020. A significant limitation of our analysis in these two reports (recognised in the reports) was that it gave no estimate of the number of asymptomatic or mildly symptomatic infections. Most of the early cases in travellers detected outside China had relatively severe symptoms, and it is likely that the majority of infections in travellers were therefore missed in border screening. We later examined the relative effectiveness of border screening in different countries [NF/038 - INQ000262587; NF/039 - INQ000236279] (24th February 2020) and estimated that two-thirds of cases exported worldwide from Wuhan remained undetected, under the still optimistic assumption that border screening in Singapore (one of the most stringent countries) was 100% sensitive at detecting infections.

121. Our estimates of the outbreak size in Wuhan were also affected by reporting delays for cases detected outside China; both in reporting newly detected international cases, and in

then updating those reports with data on the symptom onset date of each detected case. For instance, a later analysis published on 31st January 2020 by a Hong Kong group [NF/040 INQ000262601] reported data indicating that a total of 16 cases with symptom onset dates up to 19th January had been detected outside China by January 28th.

Transmissibility

122. Our second priority was to obtain estimates of the transmissibility of the novel coronavirus. Again, we undertook a relatively simple and rapid analysis making use of our estimates of the outbreak size in Wuhan and making a range of assumptions about the initial seeding of the virus into the human population (assumed to occur in early December 2019) and the generation time distribution of the virus (looking at a SARS-1-like mean of 8.4 days and a shorter MERS-CoV-like value of 6.7 days). The key conclusions from this analysis [NF/041 INQ000103222] (published on our website on January 25th but shared with the UK government and WHO two days earlier) was that any under reasonable scenario we examined, the basic reproduction number, R_0, of the new virus was above 1 (indicating self-sustaining human-to-human transmission), and most likely in the range 2-3. The more sophisticated Hong Kong University analysis published a week later and referred to above [NF/040 - INQ000262601] obtained very similar estimates.

123. I note that once we established that only a small fraction of infections in Wuhan were being detected in January 2020, it seemed likely that SARS-CoV-2 infections would vary from very severe to mild or asymptomatic. I therefore expected a substantial proportion of COVID-19 cases to show no or almost no symptoms, but for such people to still have some level of infectiousness. This was a widely held view on SAGE – the 4th February 2020 meeting concluded "Asymptomatic transmission cannot be ruled out and transmission from mildly symptomatic individuals is likely" [NF/042 - INQ000051925].

124. Hence all the COVID-19 epidemic models used by ICCRT assumed that pre-symptomatic and asymptomatic infected individuals had some level of infectiousness (albeit lower than symptomatic cases). In all the modelling of intervention policy options we undertook for SAGE between February and May 2020 (which made use of the CovidSim individual based model), one third of infections were assumed to be asymptomatic, and asymptomatic cases were assumed to be one third less infectious than symptomatic cases. Given the very limited data available at the time on the proportion of infections which were asymptomatic and their

relative infectiousness, these assumptions turned out to be reasonable. A meta-analysis of the scientific literature conducted by a US academic group and published in November 2020 concluded that 35% of COVID-19 infections were asymptomatic [NF/043 - INQ000262567], though with wide variation in estimates between studies. In early 2021, we analysed data from the REACT-1 infection survey and estimated that 55% of infections in people aged 20 or over met the definition of symptomatic used by that survey (Figure S7 in Knock et al. [NF/011 - INQ000212077]), falling to 25% in children aged under 5 – values we used in the SirCovid model from early 2021 (previously SirCovid assumed 60% were symptomatic). A systematic review and meta-analysis published in September 2020 concluded that asymptomatic individuals were 65% less likely to transmit than symptomatic cases [NF/044 INQ000262602]. We also conducted a later systematic review and meta-analysis using more studies that estimated that asymptomatic cases were one seventh as infectious as symptomatic cases, while pre-symptomatic cases were two-thirds as infectious [NF/045 - INQ000262617].

Severity

125. The IFR is the proportion of all infected people (including asymptomatic and mildly symptomatic infections) who would be expected to die from their infection. The CFR is the proportion of symptomatic cases of infection who would be expected to die. The CFR is usually easier to estimate, but estimates can vary substantially depending on the case definition used (in particular what severity of symptoms are required for someone to be counted as a case). In addition, unless one knows the proportion of infections meeting the symptomatic case definition, knowing the CFR only allows an upper bound on the potential death toll from an epidemic to be estimated.

126. IFR estimates are fundamental to determining policy strategy. A priori, making the reasonable assumption of no significant population immunity to a novel virus such as SARS-CoV-2, we would expect 80%+ of the population to be infected in an unmitigated epidemic for R_0 values of 2.5 or higher. This 80%+ figure is not sensitive to which epidemic model is used (*i.e.* all the epidemic models used in SPI-M-O would give such a number). With an IFR of 0.9-1%, 80% of the population being infected translates into approximately 500 thousand deaths in the UK. However, had the IFR been 0.1%, "only" 50 thousand deaths would have been expected – making mitigation (e.g. "flattening the curve") a more viable policy option.

127. Detailed data are essential to estimating both the CFR and IFR of a novel virus. We therefore assembled a team of up to 10 research staff and PhD students with a variety of language skills (including Mandarin) to monitor government websites, academic papers and preprints and media reports. As well as tracking case numbers by country and region, the team focussed on recording detailed clinical data regarding symptoms, delays from exposure to symptom onset, from symptom onset to case detection and hospitalisation, and from hospital admission to death. We (and I believe this was true of all SAGE participants) always kept abreast of articles published in scientific journals by Chinese groups. On 23rd January 2020, the WHO Emergency Committee reported that 4% of Chinese cases were dying (as referred to in [NF/046 - INQ000047559) – a crude estimate of the CFR. However, there were no estimates of the IFR at that time, to my knowledge.

128. On 10th February we published preliminary estimates of the symptomatic CFR and the IFR [NF/047 - INQ000228651]. This analysis was refined in the following three weeks to generate age-specific IFR estimates, with estimates of the infection hospitalisation ratio (IHR – proportion of infections requiring hospitalisation) then generated in early March. We published these IFR and IHR estimates as a preprint on 13th March [NF/048 - INQ000236303], and then in the Lancet Infectious Diseases journal [NF/049 - INQ000262584].

129. On January 30th 2020, WHO shared with MRC GIDA the detailed presentation made by China CDC to the WHO Emergency Committee on 30th January to facilitate our work in advising WHO staff on the emerging epidemic. The key information in that presentation which was not in the public domain at that time was an age breakdown of hospitalisations and deaths in China, which showed that the CFR of hospitalised cases in China increased markedly with age. I discussed this on SAGE on February 4th. Following that meeting, we used data from the presentation to generate an initial analysis of how severity varied by age and shared this with SAGE on 11th February [NF/050 - INQ000212080]. We did not make use of those data in our initial analysis of CFR and IFR presented in our Report 4 [NF/047 - INQ000228651], which did not attempt to give an age-breakdown of severity. All our early work on characterising the epidemiology of COVID-19 solely made use of publicly available data. As the WHO presentation was shared confidentially with MRC GIDA, I was not able to share the contents more widely. MRC GIDA is a WHO Collaborating Centre and we are trusted to adhere to the terms under which any data are shared, just as we do with data we receive from UK government agencies. I would also note that the UK was represented by Prof Brian McCloskey

on the WHO Emergency Committee, a past employee of PHE. The data on age-specific severity were published by China CDC on 21st February [NF/037 - INQ000215642].

130. International data sharing was referred to in the minutes of the 4th February SAGE meeting [NF/042 - INQ000051925], and there was an email exchange with Jonathan Van Tam, David O'Connor (FCDO) and SPI-M-O modellers on 7th February 2020 discussing a wish list of data to be requested from China [NF/051 - INQ000148789; NF/052 - INQ000148790; NF/053 - INQ000148791; NF/054 - INQ000148792; NF/055 - INQ000148793; NF/056 - INQ000148794; NF/057 - INQ000148795; NF/058 - INQ000148796; NF/059 - INQ000148797; NF/060 - INQ000148798; NF/061 - INQ000148799; NF/062 - INQ000148800; NF/063 - INQ000148801]. I do not recall any later discussion of that request or of data sharing by China.

131. IFR is not a static quantity – while different viruses have intrinsic differences in pathogenicity, how pathogenicity manifests as mortality depends on the characteristics of the affected population, its healthcare system and available treatments. Based on knowledge of previous severe respiratory viral infections (influenza, SARS-1 and MERS-CoV), we expected the average IFR of COVID-19 to vary with (a) the age distribution of the population affected; (b) the distribution of other clinical risk factors in the population; (c) healthcare pressures – IFR tends to increase when hospitals become close to being overwhelmed; (d) healthcare system capability (e.g. availability of intensive care beds, oxygen and ventilators).

132. We therefore anticipated that the IFR seen in the UK would differ from that seen in China, if only because China has a younger population. Equally, while we examined data from the Diamond Princess cruise ship outbreak, I doubted whether the IFR estimated [NF/064 - INQ000262586] in an affluent cruise ship client population given the best available medical care in Japan would be representative of what would be seen in a large epidemic in the UK population. In particular, due to the health requirements for being a passenger, cruise ship customers have a substantially lower prevalence (for their age) of the chronic health conditions (e.g. COPD, dementia) which are typically associated with high respiratory infection mortality. We therefore focussed on estimating IFR from data gathered in the early epidemic in China.

133. Our analysis of severity assumed that lethal cases of SARS-CoV-2 in China were being fully ascertained (since testing prioritised severely ill individuals), and (effectively) divided reported deaths by a specific date by the estimated number of infections generating those deaths,

correcting for the delay from infection to death and allowing for the growth rate of the epidemic. Our analysis built upon our extensive past experience of estimating the severity of emerging infections and knowledge of the biases such analyses are prone to.

134. The most difficult aspect of the analysis was estimating the total number of people infected in Wuhan, given that we knew mild or asymptomatic infections were missing from existing estimates. Fortuitously, between 29th January and 8th February 2020, there were 12 repatriation flights from Wuhan of non-Chinese citizens travelling back to their home country where everyone on the flight was PCR tested for infection, regardless of symptoms [NF/065 - INQ000275159; NF/066 - INQ000262615]. The proportion of passengers testing positive gave an estimate of infection prevalence (the proportion of people infected with the virus at a point in time) in the Wuhan population at the time the flights left. At the peak of the epidemic in Wuhan in late January, we estimated this to be approximately 1%. Looking at a variety of assumptions for how long an infected person would test positive on a PCR test, we were then able to estimate the total number of daily new infections occurring in Wuhan at the peak of the epidemic, and compare that number with reported confirmed cases. Assuming infected individuals test positive for 14 days, this comparison suggested that only 1 in 19 (95% uncertainty range 10-35) infections were being detected in official case counts in Wuhan at the time. This factor was then used to scale down our CFR estimates (of 18%) in China to produce estimates of the IFR in China of 0.8-0.9% in our initial analysis [NF/047 - INQ000228651] of February 10th. Over the following three weeks we considerably refined the analysis, including substantially more data, obtaining a final estimate of the IFR for the early epidemic in China of 0.7% (95% uncertainty range: 0.4%-1.3%) by early March (preprint shared with SPI-M-O on 9th March, and published on Medrxiv 13th March [NF/048 - INQ000236303]). Adjusting for the age distribution of the UK population and expected attack rate by age, we estimated this would translate into a 0.9% IFR (95% uncertainty range 0.4%-1.4%) in the UK [NF/001 - INQ000049647].

135. Our early estimates of the IFR for the UK population were later confirmed as accurate by retrospective studies of the first wave of the pandemic in the UK (prior to the introduction of dexamethasone and other treatments which improved survival) which made use of serological survey data (see section H) to estimate the total number of people infected (irrespective of symptoms) in the first wave. While there remains considerable variability between different studies' estimates of first wave IFR in the UK due to differences in data

sources, methods used and modelling assumptions, all have been broadly compatible with our early estimate of 0.9% (95% uncertainty range 0.4%-1.4%). The Imperial College based REACT antibody prevalence survey study team estimated an average IFR in England for the period March to June 2020 of 0.9% excluding care home deaths and of 1.43% including care home deaths [NF/067 - INQ000262614]. An independent joint US-France-UK academic collaboration estimated a median IFR in the first wave of 1.41% in England [NF/068 - INQ000262566]. We conducted a similar cross-country study [NF/069 - INQ000262560; NF/070 - INQ000262585] using more elaborate methods (*e.g.* accounting for seroreversion – the waning of antibodies over time) and obtained an estimate of first wave IFR (including care home deaths) in England of 1.18% without accounting for seroreversion, and of 1.07% accounting for seroreversion. The Institute of Health Metrics and Evaluation (IHME) estimated the UK IFR to be 1.57% in mid-April 2020 [NF/071 - INQ000262630].

136. While there is considerable variation in IFR estimates across different countries (even those with income levels and population age distributions similar to the UK), the IFR seen in the UK first wave was not atypical of that seen in comparable countries [NF/070 - INQ000262585; NF/071 - INQ000262630].

137. As part of the weekly real-time modelling ICCRT undertook for SPI-M-O, we have tracked changes in IFR across the first two years of the pandemic [NF/072 - INQ000262583]. IFR has changed substantially over that time, decreasing after the peak of the first wave due to improvements in treatment and reductions in hospital pressures, then increasing again in the winter of 2020 as hospitalisations increased and due to the greater severity of the Alpha variant, before decreasing once more with the rollout of vaccination in early 2021.

138. We had not attempted to estimate the infection hospitalisation ratio (IHR – the proportion of infected people who would require hospitalisation) in our pre-March 2020 research. Estimates of the risk of hospitalisation for use in UK RWC planning were initially derived from a combination of analysis of Chinese data and expert input from clinical colleagues on March 1st 2020 and then further refined over the following two weeks, as discussed below.

Adoption of severity estimates by SAGE

139. Our IFR estimates heightened concern at SAGE and SPI-M-O when discussed the week of 10th February, but some participants questioned their reliability/uncertainty, their likely applicability to the UK, and the wisdom of depending on a single group's estimates. Hence IFR

as a concept was only first mentioned in the SPI-M-O consensus statement of 17th February, but with the comment that its estimation was "difficult" and without quoting our estimate [NF/073 - INQ000074896].

140. A value of 1% was first accepted for use as an NHS "planning assumption" by SAGE on 26th February [NF/073 - INQ000074896]. It was only after the March 1st meeting discussed below that the SPI-M-O consensus statement quoted an IFR range of 0.5-1% [NF/074 - INQ000213325], following LSHTM generating their own IFR estimate (based on the Diamond Princess cruise ship data [NF/064 - INQ000262586]) of 0.5%.

141. At the SAGE meeting on 27th February when the 1% IFR value was first agreed in principle as a planning assumption [NF/073 - INQ000074896], academic groups (Imperial College, Oxford University and LSHTM) together with NHS clinical leaders and planners were tasked with organising "a working group to analyse key clinical variables for reasonable worst-case planning for the NHS". This working group meeting was hosted at Imperial College on Sunday March 1st 2020, with the following attendees (some online, most in person):

- *NHS*: Stephen Powis, Keith Willett, Andrew Jackson, Robert Shaw, Andrew Menzies-Gow, Mike Prentice

- *DHSC*: Jonathan Van Tam, Tom Irving, Paul Allen

- *University of Oxford*: Peter Horby, Mark Pritchard

- *LSHTM*: John Edmunds, Adam Kucharski

- *Imperial College*: Neil Ferguson, Azra Ghani, Christl Donnelly, Natsuko Imai, Katy Gaythorpe

142. The March 1st meeting (which lasted approximately 3h) was intended to review the IFR estimates (the age-dependence of IFR in particular – the 1% average having been previously agreed as the basis for planning at the 27th February SAGE meeting), decide on IHR values to be used for planning purposes, agree a value for the proportion of hospitalised cases that would require admission to ICU, and then agree modelled epidemic scenarios making use of those values for use for RWC and NHS planning. The group agreed to adopt the age-specific IFR estimates for the UK which ICCRT had by that time generated. Based on those estimates and clinical experience of other severe respiratory infections among the NHS and clinical experts at the meeting, attendees agreed an 8% IHR figure for the RWC. A representative epidemic profile was also agreed. The outputs from the meeting were immediately shared by NHSE staff with Simon Stevens, NHS Chief Executive, and I believe with Number 10 (though

that will need to be checked with Simon Stevens and/or Keith Willett). They were then agreed in principle as NHS planning assumptions by SAGE on 3rd March [NF/075 - INQ000119719], with the age breakdown of clinical parameters agreed on 5th March [NF/076 - INQ000074987], while the finalised RWC assumptions were agreed by SAGE on 10th March [NF/077 - INQ000262577].

143. We refined the age-specific profile of IHR over the following two weeks with further analysis of Chinese data and in discussion with NHS colleagues, and a final IHR of 4.4% was used for modelling from March 13th (including our Report 9 published on 16th March [NF/001 - INQ000049647]) until sufficient detailed data on UK hospitalisations became available in late March. The 4.4% value ended up better reflecting the 3.1 to 1 ratio of hospitalisations to deaths later seen in the first wave of the pandemic (40,800 deaths and 127,100 hospital admissions in the UK up to July 2020) than the original 8% RWC assumption.

144. We initially assumed 15% of hospitalised cases would require admission to an intensive care unit (ICU), based on expert clinical opinion gathered at the March 1st meeting. This figure was later revised upward to 30% by March 14th based on clinical reports from Italy and discussions with NHSE colleagues about the emerging clinical picture being seen in the UK. Report 9 therefore assumed a 4.4% IHR and that 30% of hospitalised patients would require ICU. These parameter estimates were then further refined over the next 2 months using individual-level hospital data collected by the CHESS and CO-CIN systems (see section H).

145. It was recognised by SAGE, SPI-M-O and NHSE that the IFR value of 1% assumed a functioning healthcare system, and that the IFR would likely exceed this value were hospitals so over-subscribed that treatment was unavailable to a substantial proportion of those requiring it – potentially reaching close to the IHR. However, we did not inflate IFR values in Reports 9 [NF/001 - INQ000049647] and 12 [NF/002 - INQ000262593] to account for this effect, given that to do so would have required further speculative assumptions to be made and that the numbers of deaths for the scenarios which exceeded healthcare demand were already large. Modelling undertaken by Steven Riley arguing against the feasibility of mitigation strategies included substantial healthcare capacity related inflation of IFR [NF/078 - INQ000222002] (see section G), and some healthcare system related scaling of mortality was included in some of our later global analyses [NF/025 - INQ000262629].

146. While understandable given the magnitude of decisions needing to be made, the three-week delay between the availability of our first IFR estimates and the initial proposal of a 1% IFR value for policy planning (26th February) and the further week delay until final sign-off (3rd–10th March) was frustrating for me at the time and clearly regrettable in retrospect. It is difficult for me to establish a clear cause, other than the caution felt by SAGE in accepting estimates with a high level of uncertainty until that uncertainty (slightly) declined. The further delay until broad government acceptance that the RWCS was a *likely* rather than just a "worst case" scenario was even more regrettable.

147. While I don't have a detailed understanding of how IFR estimates were being perceived within government at that time, the labelling of a 1% IFR value as a "reasonable worst case" (rather than as the central scenario) may have contributed to a sense that reality was unlikely to be as bad as our modelling suggested. It is plausible that this may have contributed to delays in formulating and enacting policy responses, and in communicating the likely severity of the unfolding epidemic to the public. It was only at the March 13th SAGE meeting [NF/079 - INQ000109142] that NHS representatives stated on the record that NHS capacity would be overwhelmed many times over under any of the mitigation policy scenarios we had been modelling up to that time. I return to these issues in section G below.

G. Pre 23rd March 2020 modelling of non-pharmaceutical interventions (NPIs)

148. The decision by the UK and many other countries in March 2020 to adopt long-term use of NPIs to suppress (*i.e.* reduce R to below 1) COVID-19 transmission until vaccines were available was a paradigm shift in the global response to a pandemic. Such an approach was never anticipated in prior UK pandemic preparedness planning, which at most had considered transient NPI use to mitigate (but not stop) a pandemic wave. Here I briefly review modelling of the use of NPIs prior to the pandemic and its use in preparedness planning, before chronologically reviewing the work of ICCRT on modelling NPI strategies during the first wave of the UK pandemic. In doing so, I try to give a sense of how my thinking and that of SAGE evolved.

Background

149. I note that ICCRT could not have delivered more work (both modelling of interventions and the analysis discussed in section F above) than we did in January-March inclusive – I was working 18+ hour days, 7 days a week that entire period, as were many of my team.

150. At the outset I should define two terms:

a. *Suppression* – the use of stringent interventions to drive R below 1 considerably earlier than would occur otherwise (where epidemics turn over due to herd immunity). The "containment" strategies I examined in pre-pandemic modelling (see below) were extreme examples of suppression (where incidence of infection is driven to zero), though the term "containment" has also been used in a wide variety of other contexts.

b. *Mitigation* – the use of interventions to reduce transmission, but not to the extent that R is driven below 1 in the short-term. In a mitigated epidemic, the virus would continue to spread but the peak of the epidemic wave would be reduced and the time-period of the epidemic extended compared with not intervening. Mitigation can achieve moderate reductions in the proportion of the population infected in an epidemic, and thus in the expected numbers of hospitalisations and deaths.

151. I have 20 years of experience researching interventions to control the spread of respiratory and other viruses. This includes work on SARS-1, 1918 "Spanish" Influenza, H5N1 "bird flu" and related preparedness research, Ebola, MERS-CoV and, most recently, COVID-19.

152. Interventions deployed in an epidemic fall into three broad categories:

a. Vaccines – to prevent infection or illness.

b. Therapeutics (e.g. antiviral drugs) – to treat illness and/or reduce infectiousness.

c. Non-pharmaceutical interventions (NPIs) – measures to reduce infectious contacts between infected and uninfected individuals.

153. NPIs span a wide range of measures which can also be divided into three classes by the population groups targeted:

c. *Border controls* – targeting people entering a region (usually a country). Testing and quarantine of international visitors was introduced by most countries during the COVID-19 pandemic, with varying levels of effectiveness.

d. *Case-focussed measures* – targeting suspected or confirmed cases of infection. Isolation of suspected or confirmed cases (in quarantine units or at home) is the most obvious measure and was introduced by most countries during the pandemic. The next step is to isolate contacts of cases – identified either through group membership (*e.g.* members of the same household, school class or workplace group) or via explicit contact tracing. Most countries introduced such measures, again with varying levels of effectiveness.

e. *Community-focussed measures* – these measures aim to reduce (infectious) contacts between all individuals, on the basis that not all infected individuals can necessarily be identified rapidly enough for purely case-focussed measures to achieve sufficient control of transmission. Most of the social-distancing measures adopted by different countries during the pandemic fall into this category (*e.g.* working from home, stay at home orders, closing schools and universities, closing hospitality venues, closing non-essential retail, limiting gatherings of people from different households, recommending minimum physical separation distances), as do mask-wearing mandates.

154. Randomised clinical trials (RCTs) represent the gold standard for gaining evidence about the effectiveness of interventions. However, RCT results are lacking for nearly all NPIs, due to the difficulties of running such studies and the relatively rare situations such interventions are deployed. Face masks are the one exception, but prior to the pandemic only one RCT of face mask use in a community setting had been conducted, to my knowledge (all others being in clinical settings). I had worked with an Australian clinical colleague on that one study, which examined the effectiveness of mask use at preventing influenza transmission in households.

155. Despite the lack of RCT evidence, considerable effort was invested in the analysis of observational datasets to assess the likely effectiveness of NPIs as part of pandemic planning following the 2003 SARS-1 outbreak and the re-emergence of H5N1 in South East Asia in 2005. I was heavily involved in much of this work via my then involvement in the US NIH-funded MIDAS network.

156. My research in the period 2004-2008 had two streams:

a. Estimating the effectiveness of NPIs from historical data. This spanned evaluation of the effectiveness of border measures, analysing data from the 1918 influenza pandemic on

mortality and the timing of NPI use in different US cities to estimate the effectiveness of the NPIs (mostly social distancing measures such as closure of schools, bars and churches) used then, and analysis of seasonal influenza data from France to estimate the potential impact of school closure as an NPI to limit pandemic influenza transmission.

b. Using simulation modelling to examine different strategies for NPI, antiviral and pre-pandemic vaccine use in a future lethal influenza pandemic. I first examined the feasibility of containing (*i.e.* eliminating) a nascent pandemic in its source location – which concluded elimination was only likely feasible if the outbreak was detected at a very early stage and intensive NPIs were deployed. I then examined the layered use of NPIs (border-, case- and community-focussed) and stockpiled antivirals to mitigate the health impacts of an influenza pandemic. This work informed pandemic planning in the US and UK (*e.g.* page 41 of reference) . The research made use of the large-scale individual-based simulation model which was later adapted to model NPI strategies for COVID-19 [NF/001 - INQ000049647].

157. Given the lethality of a future pandemic virus is unknown, I did not model healthcare demand or mortality when simulating potential future pandemic influenza scenarios in the work described above. My analysis solely focussed on the impact of interventions on the potential total number of infections or symptomatic cases.

158. In UK pandemic preparedness planning up to 2018, DHSC analysts (at that time led by Dr Peter Grove) translated symptomatic case numbers into predicted health-care demand and mortality using agreed UK government RWC assumptions for an influenza pandemic. I believe there was considerable work examining the implications of these planning scenarios for the NHS (*e.g.* emergency triage protocols, resource implications such as refrigeration capacity, body bags, PPE), but I was not personally involved in that work.

159. Thus, the major modifications made to the simulation code (later named CovidSim) in January-March 2020 to model potential trajectories of the COVID-19 pandemic was the development of a detailed representation of healthcare demand (*i.e.* age-varying probabilities of admission and length of hospital stay in general wards and intensive care units).

160. Richard Hatchett, now CEO of CEPI, but then a member of the White House Homeland Security team can give a policy-focussed perspective on much of the US pandemic preparedness planning regarding NPI use undertaken between 2005 and 2009.

161. The past work on pandemic mitigation described above did not model long-term large-scale use of intense community-focussed NPIs to suppress (*i.e.* achieve R<1) influenza transmission for many months; rather, the focus was on the extent to which "feasible" NPIs (case isolation, household quarantine, time-limited closure of schools only after cases are detected in them, limited reactive closure of a minority workplaces with outbreaks, border restrictions) together with antiviral use might mitigate a pandemic wave – *i.e.* reduce numbers infected and/or flatten and delay the peak of the pandemic in an affected country.

162. I cannot recall any discussion of the potential use of long-term suppression policies (outside the containment at source context) in meetings with public health policymakers in either the UK or US prior to 2020. The goal of the pandemic preparedness planning I was involved in prior to 2020 was to minimise the health impact of a pandemic while still allowing society to function as close to normally as possible, thus limiting economic and social disruption. Containment/suppression was only viewed as a viable option for outbreaks of high severity pathogens which were limited in size and geographic scope and where the elimination of the virus from the human population therefore seemed feasible. Real-life examples include SARS-1 and the West Africa and DRC Ebola epidemics – where control achieved elimination of those epidemics largely via case-focussed measures (case isolation and contact tracing).

163. Throughout the COVID-19 pandemic, I was rarely involved in advising on the operational details of policy implementation. One exception, via NERVTAG, was the recommended duration of case isolation and household quarantine (see para 214). This discussion focussed on the trade-offs between the effectiveness and practicality/cost of different durations of case and household isolation.

164. Trade-offs between potential impact (on transmission or mortality) and cost (economic or social) exist for nearly all interventions. A related trade-off is between effectiveness and practicality; a less onerous or more practical intervention may achieve a higher impact than a more onerous one which is poorly adhered to, or an intervention where the practical challenges of implementation impair effectiveness. Rarely can one rule out *some* level of effectiveness for a potential intervention, so frequently the issue facing policymakers is whether the cost and/or operational challenges of implementing a measure is *appropriate* given its likely effect. Border controls are one example; in my view, the border measures adopted by the UK at various times during the pandemic had little impact on the final

mortality and healthcare burden the country experienced. However, throughout the pandemic I was conscious that policymakers had to balance public and political pressures to implement border controls against implementation challenges and the economic and social impacts of such measures. Mask mandates are another example; SAGE eventually recommended mask use for indoor public spaces in April 2020 not because the expected impact was high (there was limited evidence to support a major impact), but because as a relatively low-cost (socially and economically) intervention, the benefits of mask use (even if limited) outweighed the costs.

Before March 5th 2020

165. The UK's published strategy for influenza pandemics involves three phases: detection/assessment, treatment and recovery. The treatment phase presumes availability of stockpiles of effective antivirals.

166. This three phase plan was replaced by a four-phase strategy in the COVID-19 UK Action Plan on March 3rd 2020: Contain, Delay, Research, Mitigate [NF/080 - INQ000280727]. While individual elements of the Action Plan were discussed by SAGE and NERVTAG prior to its publication, the overall framing of the Plan (*i.e.* its four phases) and its text were not reviewed by SAGE or NERVTAG.

167. I thought the stated goals and ordering of the Action Plan phases were broadly sensible, other than that "research" is not a phase, but a continuous activity undertaken throughout a pandemic, However, the policies in place at the time the Plan was announced (travel advisories/restrictions, enhanced surveillance of travellers returning from identified high risk regions, limited contact tracing around index cases) had little chance of preventing establishment of infection in the UK (the goal of "Contain") and were unlikely to delay establishment by more than a few days. I and John Edmunds made this point repeatedly on NERVTAG and SAGE from late January onwards, backed by the research our teams were undertaking at ICCRT and LSHTM [NF/039 - INQ000236279]; [NF/038 - INQ000262587; NF/081 - INQ000212209]. Indeed, my working assumption by the start of March 2020 was that community transmission was already underway in the UK. I note that the border measures adopted by the UK between January and March 2020 seemed to be drawn from the same play book as the early UK response to the 2009 H1N1 influenza pandemic, with no greater success in 2020 than in 2009.

168. The limited effectiveness of the border measures adopted by the UK (and discussed extensively by NERVTAG in January 2020 [NF/082 - INQ000047820]) was because those measures were always a balance between what might be effective (namely, very intensive measures such as extensive travel bans and passenger testing) and what was viewed by government as practical, given personnel and testing constraints. This was well appreciated by the CMO and GCSA; Chris Whitty emailed Patrick Vallance, Jonathan Van tam, Charlotte Watts, John Ashton, John Edmunds and me on February 2nd 2020 to discuss his provisional view on the likely effectiveness of border measures/travel restrictions [NF/383 - INQ000047654]. Both John Edmunds and I largely agreed with his relatively pessimistic analysis that UK-only China-focussed measures would likely only achieve minor delays in slowing (UK) transmission, but that impacts would be greater if multiple countries took concerted action. I also gave some approximate assessments of the timescale of delays different options might – suggesting that a combined UK/EU cessation of travel to/from China might achieve a three-week delay.

169. Before March 2020, UK COVID-19 surveillance was focussed on symptomatic cases in travellers coming into the UK meeting a case definition which incorporated clinical (i.e. types of symptoms) and geographic (recent travel from a "high risk" area) aspects. This case definition changed frequently between January and March as the pandemic spread to additional countries in Asia and then Europe, and as more data on symptoms became available. However, the definition always lagged the pandemic's geographic spread, and was nearly always focussed on clearly symptomatic cases. The exceptions to the requirement for symptoms to be present were the 11th February move to compulsory 14 day self-isolation for all travellers returning from Hubei province in China, and the 26th February extension of this to travellers from Iran, Northern Italy and part of South Korea. Except in a very few instances (e.g. repatriation flights from Wuhan) the UK did not use quarantine facilities for self-isolation, but recommended isolation at home. I further discuss the limitations of early COVID-19 testing and surveillance in the UK in section H.

170. On January 29th, Chris Whitty emailed John Edmunds and me to ask about intervention options for delaying a pandemic wave in the UK [NF/083 - INQ000148970]; [NF/084 - INQ000148971]; [NF/085 - INQ000148972]; [NF/086 - INQ000148974]. He stated that "modelling what we could do to delay the upswing of an epidemic would be really useful". In my reply I stated that "delaying arrival requires either stopping travel from China or very intensive

screening and follow-up of travellers". I then talked about how case isolation/contact tracing and/or community-level interventions might be employed to delaying the peak of any UK epidemic. I noted that ICCRT were then gearing up to be able to undertake such analysis but warned that given the many uncertainties about the epidemiology at that time (*e.g.* around the age distribution of infections and symptomatic disease, a point John Edmunds made), any modelling undertaken would largely be restricted to scenario analysis, with some potential risks of misinterpretation.

171. The February 3rd SAGE meeting clearly set out the limits of the likely effectiveness of border measures and travel restrictions [NF/087 - INQ000051883]. The 4th February SAGE meeting noted that perhaps 1 in 15 cases were being detected in China at that time, and perhaps 1 in 4 cases outside China [NF/042 - INQ000051925]. The 1 in 4 figure (which in retrospect was highly optimistic) inevitably meant that UK border measures were going to have very limited effectiveness.

172. Other than some discussion of the potential effects of the policies enacted by China on 23rd January, community focussed-NPIs were first discussed by SAGE in a UK context on 13th February [NF/088 - INQ000106109], drawing on a SPI-M-O consensus paper on NPIs of 3rd February [NF/089 - INQ000051882] and a second SPI-M-O consensus paper on the potential impacts of mass school closure of 10th February [NF/090 - INQ000075404]. The only community-focussed measures discussed in those consensus documents were school and university closure, restricting mass gatherings, and mask wearing. The framing of the SPI-M-O documents and of those discussions at SAGE were solely around the use of NPIs to delay the peak of an epidemic, not around suppression (achieving R<1).

173. Given the relatively high estimates of R_0 and the limited data available, I initially had some scepticism that the measures adopted in China from 23rd January would actually manage to reduce infectious contact rates sufficiently to drive R to below 1. However, by the third week of February, the trends in the Chinese case data (both nationally and at province level) were sufficiently clear that I was persuaded that intensive NPIs were able to achieve suppression of COVID-19 transmission. I was convinced of the feasibility of implementing such measures in a UK context after they were adopted in Northern Italy in late February.

174. Community-focussed NPIs were next discussed by SAGE on February 20th, when a SPI-M-O paper on the potential impacts of school closure on pandemic peak height and timing was

discussed. This paper included an annexe written by ICCRT [NF/091 - INQ000075775] presenting the potential impacts of school closure under a range of epidemiological scenarios spanning children being as susceptible and infectious as adults to being substantially less likely to exhibit symptoms and being somewhat less likely to be infected than adults. SAGE concluded "it is possible that school closures could have a modest impact on delaying the peak of an epidemic". This was the first time that modelling making use of the CovidSim individual-based model was tabled at SAGE. Discussion of other measures to limit spread was postponed pending better understanding of the likely sensitivity of surveillance systems then being proposed and of the objectives of NPI use in light of the "key challenges the NHS is seeking to mitigate".

175. We accelerated our work on modelling NPIs in response to both the discussion at the February 20th SAGE meeting and the news on February 21st of a cluster of locally transmitted cases in Lombardy Italy [NF/092 - INQ000262575] which resulted in authorities there introducing a stay-at-home order. I had been concerned for the previous week that the UK needed to urgently formulate a strategy for how it would respond to the first indications of local transmission within the country, or risk being "bounced" into a similar response to Italy, which at the time I felt would be an unsustainable (and certainly economically very costly) policy. I emailed Chris Whitty, Patrick Vallance and a number of other SAGE participants to that effect on 21st February [NF/093 - INQ000148969]. By that time, I felt it was a matter of when rather than if local community transmission would be discovered in the UK, and thus that the UK should move away from its stated "containment" strategy (*i.e.* focussing on border measures, which in my view were ineffective) to an evidence-based mitigation strategy (*i.e.* targeted, layered use of NPIs to reduce transmission). Chris Whitty disagreed regarding the inevitability of local spread and felt it was premature to talk about ending containment. Overall, I felt there was insufficient urgency in government planning at that time, a view shared by John Edmunds and Jeremy Farrar. I return to this point further below.

176. ICCRT therefore prepared three papers for the February 25th SAGE meeting: two on surveillance system design and sensitivity [NF/094 - INQ000262616], (discussed in section H) and one on the effect of community-focussed NPIs [NF/095 - INQ000236282].

177. Our 25th February paper on NPIs used the individual-based simulation model CovidSim (see section E for details) later used in our Report 9 to examine the potential impacts of a range of combinations of different community-focussed NPIs. The paper considered (a)

closure of schools and universities; (b) home isolation of cases for 7 days; (c) home isolation of other members of the household of index cases for 14 days; (d) mass social distancing consisting of a 75% reduction of all interpersonal contacts other than in the home, school, university or workplace and a 25% reduction in workplace contacts (household contacts were assumed to increase by 25% in compensation for reduced contacts elsewhere). Policies were assumed to be imposed for 6 months. The paper was caveated with the sentence "With the exception of school closure, we do not have reliable estimates of the impact of these policies, even for influenza. Plausible values have been selected".

178. All results in that paper were presented in terms of impacts on overall symptomatic case incidence; at the time of preparation, incorporation of hospital demand and mortality (and the age-specific severity estimates underpinning those) had not been finalised. However, I believe that SAGE participants were aware of the likely scale of mortality the scenarios represented if the IFR was anything close to 1%.

179. Our paper presented epidemic scenarios over a one-year timescale. The CovidSim model (in common with all SPI-M-O models prior to the summer of 2020) assumed COVID-19 infection would induce immunity which would protect against reinfection over that timescale, based on what was known about seasonal, endemic human coronaviruses. This assumption turned out to be largely justified by retrospective studies of immunity following COVID-19 infections, at least until the emergence of the Omicron variant which escaped immunity from both vaccines and past infection to a substantially larger extent than prior variants (see section K). In addition, there is now good evidence that immunity induced by COVID-19 infection, like that induced by vaccines, is higher and more sustained than that against infection, even in the face of the Omicron variant. In reality, reinfection with COVID-19 only became a significant factor in the epidemiology of the pandemic in the UK from December 2021, with the advent of Omicron.

180. The 25th February paper noted that "Combination policies are predicted to be sufficiently effective at reducing transmission to give rise to double-peaked epidemics (second peak in late 2020) when the interventions are lifted" and that "Aggressive NPIs may have a substantial impact on COVID-19 transmission, potentially dramatically slowing epidemic growth or reducing R to below 1 while in operation", but commented that "However, the primary impact of such measures is to delay transmission and reduce peak incidence; when they are lifted, transmission can be expected to resume given the measures only protect the population while

in operation (unlike vaccination)". It further commented that "measures which are too effective merely push all transmission to the period after they are lifted, giving a delay but no substantial reduction in either peak incidence or overall attack rate".

181. Thus, the framing of that paper followed that of all prior discussions (both pre-pandemic and up to that point in the pandemic) of the goals of NPI use (whether border-, case- or community-focussed): namely to delay and reduce the peak of the epidemic and to achieve modest reductions in the total number of people infected. While the scenarios modelled in that paper included what would later be termed suppression strategies, at that time I did not foresee that the adoption of such strategies would be seen as the "least worst" policy option less than three weeks later.

182. Indeed, in the SAGE discussion of our paper on February 25th, Chris Whitty's immediate response to seeing the more aggressive NPI scenarios was that pushing the epidemic to the autumn or winter might worsen the consequences of the pandemic, given the higher likely incidence of other respiratory infections and the rising seasonal pressures on the NHS. Thus, the NPI modelling requested by SAGE for the following two weeks focussed on "flattening the curve" policy options which could avoid a second peak of transmission once NPIs were lifted.

183. The SAGE minutes covering this discussion are informative of the then thinking [NF/096 - INQ000087503]. To precis those minutes: the modelling is highly uncertain and will likely remain so until after decisions need to be made, but both it and emerging evidence from Singapore and China suggest that NPIs can push R down to 1 or below, that such an impact could be realistically expected in the UK, and that "NHS needs must be considered in any decisions to alter the epidemic curve".

184. On February 27th the SPI-M-O secretariat emailed a request from Jonathan Van Tam for views on "the impact of holding Six-nations Rugby matches" on the risk of importing cases or the risk of spreading COVID-19 [NF/097 - INQ000148982; NF/098 - INQ000148983; NF/099 - INQ000148984; NF/100 - INQ000148985; NF/101 - INQ000148986; NF/102 - INQ000148987; NF/103 - INQ000148988; NF/104 - INQ000148989; NF/105 - INQ000148990; NF/106 - INQ000148991]. On the former issue, I responded that "*I now believe it is >95% certain that transmission is already established here. From that perspective, holding the 6 Nations matches will make no difference. But if HMG wants to adopt the optimistic assumption that transmission has not established itself here, then the question to ask is how much will the 6*

Nations matches increase travel to the UK from Europe. I suspect not by much, given the large daily volumes of travellers anyway". On the second question, I felt that the risk of transmission in a stadium was likely no greater (and perhaps less) than that associated with watching the match in a pub. There was broad consensus among the SPI-M-O members who responded that the risks were low but not zero. The topic was then discussed at the SAGE meeting later that day, and similar conclusions were reached [NF/107 - INQ000106129].

185. The 27th February SAGE meeting [NF/107 - INQ000106129] reviewed a SPI-M-O consensus statement on the potential impact of different NPIs [NF/108 - INQ000075403] – the formulation of this document involved input from multiple modelling groups and its conclusions on the effect size of individual interventions largely agreed with our earlier paper[NF/095 - INQ000236282].

186. Also at that meeting, the co-chairs presented a set of current priorities for SAGE [NF/109 - INQ000248850]. In my view those priorities accurately identified the then key evidence needs and substantially shaped SAGE discussions over the following 3 weeks.

187. As detailed in section F, the 3rd March SAGE meeting [NF/075 - INQ000119719] agreed in principle to use the outputs of the March 1st SAGE working group meeting (hosted at Imperial College) as the basis for RWCS NHS planning (though final sign-off occurred in the following two SAGE meetings).

188. The 3rd March meeting was also the first where those planning assumptions (even if viewed as preliminary at that time) were used in the modelling of NPIs to project potential hospital demand and mortality. The meeting considered modelling of mitigation strategies undertaken independently by ICCRT [NF/110 - INQ000195872] and LSHTM [NF/111 - INQ000195871].

189. The 3rd March meeting was also the first time that inclusion of "shielding" was included in our modelling [NF/110 - INQ000195872] – represented as a reduction in contact rates of those over 65 years of age, rather than the entire population. I believe the concept of shielding was first proposed on SAGE by Graham Medley in one of the last two meetings in February, but I don't remember which and the topic wasn't minuted.

190. I had sent an initial version of the modelling of shielding of over-65s to Chris Whitty and Patrick Vallance on 29th February [NF/112 - INQ000148992; NF/113 - INQ000148993; NF/114- - INQ000148994; NF/115 - INQ000148995; NF/116 - INQ000148996; NF/117 - INQ000148997;

NF/118 - INQ000148998; NF/119 - INQ000148999; NF/120 - INQ000149000; NF/121 - INQ000149001]. In the same email exchange, Chris Whitty asked me to try and look at the effect of restricting mass gatherings.

191. Hence for the modelling [NF/110 - INQ000195872] considered by the 3rd March SAGE meeting (a preliminary version of which I sent to the same email exchange [NF/112 - INQ000148992; NF/113 - INQ000148993; NF/114- - INQ000148994; NF/115 - INQ000148995; NF/116 - INQ000148996; NF/117 - INQ000148997; NF/118 - INQ000148998; NF/119 - INQ000148999; NF/120 - INQ000149000; NF/121 - INQ000149001]), I modelled the cessation of mass gatherings as a 16% reduction in transmission occurring outside the home, school or workplace contexts. The paper took a very broad view of mass gatherings – defined to include all hospitality venues. It stated "Includes shutting, in order of significance, bars/pubs, restaurants, cinemas, night clubs, sporting fixtures, places of worship and theatres. These represent about 12m contact hours of activity per day, or 5.3% of all hours outside home, school or work. Assuming a 3-fold higher risk of transmission than other activities, preventing them might reduce transmission outside household, school or work contacts by 16%". I based the contact-hour assessment on a very rapid search for time-use studies (I didn't retain records of what sources I found, unfortunately). The three-fold high transmission risk was a best-judgement assumption of the likely effects of crowding. Some support for the net 16% figure was later provided by a PHE analysis of time-use data [NF/122 - INQ000120644].

192. The 3rd March SAGE meeting also considered a rapid review of what NPIs had been adopted by different countries in the pandemic thus far, and what we could say about their effectiveness [NF/123 - INQ000195870]. This had been conducted over the previous few days by a joint LSHTM-ICCRT team. It noted that data on the effectiveness of NPIs for COVID-19 were currently sparse, but that "Timely implementation of control measures will have a greater impact on the COVID-19 epidemic, but the early lifting of control measures could lead to an increase in case numbers, as shown for influenza control".

193. John Edmunds and I both emphasised in late February that infection control around social care would be vital for shielding to be effective, given the most vulnerable elderly are dependent on social care provision (home visits or via residential care homes) and have a high frequency of contact with the health system. While SAGE minutes first mention potential risks around social care for the March 10th meeting [NF/124 - INQ000109125], my recollection is that care homes had been discussed as a concern in late February following reports of the

first large-scale COVID-19 outbreak in a US nursing home and discussion of the concept of shielding of the elderly. Indeed, the earliest occasion may well have been earlier – unfortunately I do not remember every discussion at SAGE. However, in the email exchange of 29th February-1st March [NF/112 - INQ000148992; NF/113 - INQ000148993; NF/114- - INQ000148994; NF/115 - INQ000148995; NF/116 - INQ000148996; NF/117 - INQ000148997; NF/118 - INQ000148998; NF/119 - INQ000148999; NF/120 - INQ000149000; NF/121 - INQ000149001; NF/125 - **INQ000087584**; NF/126 - INQ000056158] referred to above, I noted that achieving the policy effectiveness for shielding that we had assumed "would require intense infection control around care homes...", suggesting that care homes had been a recent topic of discussion.

194. No SPI-M-O models (including ours) used up to 23rd March explicitly modelled care homes or within-hospital transmission. This was largely due to lack of data (on the demography of the care home sector and on likely transmission rates in both care homes and hospitals) and lack of time to make relatively substantial changes to models. We incorporated care-homes in our UK models (SirCovid and CovidSim) from May 2020. We later undertook retrospective analyses of the extent to which improved infection control might have reduced mortality in care homes in the first wave of the pandemic [NF/011 - INQ000212077].

195. The ICCRT and LSHTM papers on the impact of NPI-based mitigation strategies presented on 3rd March reached qualitatively and quantitatively similar conclusions. A SAGE/GoS summary paper [NF/125 - **INQ000087584**; NF/126 - INQ000056158] was also considered, which noted that "whatever the reduction in peak NHS bed demand achieved by these interventions, in the reasonable worst-case scenario demand will still greatly exceed supply".

196. However, the meeting minutes noted "SAGE advised that infection attack rate and infection fatality rate are likely to be lower than the reasonable worst case, but this will depend on the effectiveness of potential interventions covered above". As I have stated above (section F), I did not personally agree that the IFR was likely to be significantly lower than the RWC agreed figure of 1% and stated as much at a number of SAGE meetings. In retrospect, a number of studies later estimated the IFR seen in the first wave of the UK pandemic was approximately 1%, as discussed in section F above.

197. The 3rd March SAGE meeting[NF/127 - INQ000106152] involved a substantial discussion on what the optimal timing for NPI-based mitigation strategies would be, leading to a

commission to SPI-M-O on this topic. We responded to this request by undertaking more refined modelling of mitigation strategies, specifically examining whether local authority level healthcare-demand driven thresholds for initiating NPIs could be more effective (further reduce peak healthcare demand) or efficient (measures could be in place for a shorter time), examining a wide range of policy options (including packages of interventions which would suppress the first wave of transmission), thresholds for initiating policies and two policy durations (2 and 3 months).

198. This work was presented at the 5th March SAGE meeting [NF/128 - INQ000221998], together with a modelling paper from University of Edinburgh [NF/129 - INQ000103519] on the same topic, and a summary paper from SPI-M-O which also included modelling from University of Cambridge [NF/130 - INQ000262574].

199. As discussed in more detail in Section I, the timing considerations for mitigation strategies are fundamentally different than those for suppression; for the former, interventions need to be tuned to give maximum reduction of the peak of the epidemic while not achieving suppression. They can therefore be introduced at a later stage in the epidemic (preferably a month or so before the peak) than suppression policies – where it is always optimal to act as early as possible. All the discussion of intervention timing at SAGE prior to March 13th was in the context of a mitigation strategy.

March 5th-March 15th 2020

200. The 5th March SAGE meeting was significant in making relatively clear policy recommendations, all implicitly with the overarching constraint that the policy goal was mitigation rather than suppression. At that meeting, an animated discussion lay behind the summary statement "There are currently no scientific grounds to move away from containment efforts in the UK", a statement I did not personally agree with. My concerns were because that was also the first SAGE meeting which acknowledged that community transmission was underway in the UK (detected as soon as limited ICU and GP surveillance had been started a few days earlier – see section H). I (quite possibly imperfectly) recall that Chris Whitty was uncomfortable with the idea that the UK would be the first European country to abandon containment and that to do so might (in light of the Action Plan published two

days earlier) necessitate an immediate switch to community NPIs as envisaged for the "delay" stage.

201. Building on the earlier discussions referred to above, the 5th March SAGE meeting also revisited large gatherings and concluded that "there is no evidence to suggest that banning very large gatherings would reduce transmission". The reasoning for this conclusion was outlined well in a later SPI-M-O paper [NF/131 - INQ000195883]. An additional consideration discussed on 5th March was that banning public attendance at, say, football matches might just lead to compensatory levels of population mixing in social venues such as pubs.

202. The unease I felt about the pace of government planning which motivated my 21st February email to Chris Whitty and Patrick Vallance (see para 175) had grown in the subsequent two weeks and I believe was shared by John Edmunds and Jeremy Farrar. My specific concerns were that:

a. From late February, when transmission in the UK was first documented (and the first cases completely unlinked to travel clusters were detected), I was convinced there was substantial community transmission of COVID-19 occurring within the UK which was not being detected, due to the lack of sensitive systematic surveillance and the criteria being used for most testing (namely needing a travel history to high risk countries). In my mind, it was just the extent of transmission which remained uncertain. I was therefore concerned that the "contain" phase was continuing and that no NPI strategy for the "delay" phase had been finalised.

b. I was also very concerned at the delays in getting a sufficiently sensitive surveillance system up and running to allow us to accurately assess the stage the epidemic had really reached in the UK. This was a frustration shared by several members of SAGE, including Patrick Vallance, and is discussed in more detail in section H below.

c. From late February, I had a high degree of confidence that the unfolding epidemic would impose a very high health burden (in terms of deaths and hospitalisations) in the UK which would lead to hospital demand several-fold higher than the "normal" NHS bed or ICU capacity. This conclusion held even if the IFR was at the lower end of the uncertainty range we had estimated (*i.e.* about 0.4%). I was concerned that it was unclear to me whether the government had really accepted this reality and therefore if sufficient resources and priority were being given to the required NHS (and cross-government) planning. If

mitigation was to be used as a realistic policy option, I would have expected to see a war-time emergency effort to build field hospitals, and plan for (likely brutal) emergency triage procedures and associated high mortality. While I did not have good visibility of operational planning (that not being the role of SAGE), it did not appear that such planning really accelerated to the pace needed until after 13th March.

d. A related concern was that messaging by the government didn't sufficiently clearly communicate the scale of the threat publicly, at least until close to 16th March. There appeared to be significant concern about appearing alarmist given the scientific uncertainties (see also the end of section N).

203. However, I was conscious that the role of SAGE was to offer scientific advice, not operational or policy advice (this was repeated to SAGE participants on multiple occasions), and also that I was not aware of most of the operational planning going on within government. Hence while I was by no means silent, I voiced the concerns listed above only in measured terms until the week of 9th March 2020, focussing on repeating my view that a 1% IFR was a likely value (not just a RWC), trying to convince officials that we really were likely to see hundreds of thousands of deaths, and that testing and systematic surveillance needed to be scaled up urgently.

204. I did not strongly advise for a switch to a suppression strategy prior to March 13th, in part because of my belief that it isn't the role of scientific advisors to determine policy (particularly such consequential policy), but also because I was very conscious of the huge economic and social costs which would be entailed by long-term and intensive use of NPIs, and was unsure that such measures were sustainable for the extended period that they would be required (namely until vaccination became available).

205. In that context, it is notable that I don't recall any serious discussion at SAGE of the likely timescale over which vaccines might become available until late March or early April 2020.

206. I expressed my feeling that policy officials seemed to be viewing our (and LSHTM) modelling projections as if they were unlikely to happen when I directly addressed the Number 10 officials present at the SAGE meeting on 10th March [NF/124 - INQ000109125] and asked them if they knew what an epidemic with 4000-6000 deaths per day would feel like. I then followed up later that day with an email to Ben Warner at Number 10 [NF/132 - INQ000149008; NF/133 - INQ000149010] which stated:

So long as the PM and Cabinet accept and understand this is what is likely to happen and are still happy to proceed with current plans, then there is a rational basis to that decision which I would say the science supports. I might suggest the messaging start quickly changing to reflect the likely magnitude of event though. This event is in the natural disaster category, and the cure (e.g. massive social distancing, shutdowns) could be worse than the disease. Nor do we know whether we could sustain very intensive China/Italy style policies for the many months required before a vaccine is available.

But what would be the worst outcome – in my opinion – would be to go for mitigation (the policy package currently being discussed) and for the health, social and political cost to be judged later to be unacceptable – necessitating a policy pivot in the midst of what will already be a national crisis. If the numbers on those graphs (for the mitigated epidemic) are viewed as unacceptable, that message needs to come down from the top now.

207. A paper written by my MRC GIDA colleague Steven Riley was considered at the 11th March SPI-M-O meeting and the 16th March SAGE meeting [NF/078 - INQ000222002]. It advocated for suppression to be adopted, arguing that mitigation would likely fail in the goal of getting the epidemic "over with" in a few months. The paper hypothesised that in the face of a lethal pandemic which overwhelmed the NHS, spontaneous behaviour change by the population would lead to partial suppression and the epidemic being extended in duration to as much as 18 months, with 400 thousand to 1.7m lives being lost. The model assumed a 4-5 fold increase in IFR in situations where healthcare demand was overwhelmed.

208. I had discussed the ideas behind his paper, the approach it took and his wider views on the pandemic several times with Steven from late February. We took different perspectives on the extent to which scientists should advocate for specific policy options. Furthermore, while I agreed that incorporating risk-related spontaneous behaviour change into models is an important (if difficult) research priority, I thought his hypothesis that such effects would extend an epidemic by 12 or more months was not supported by the historical record. I also agreed it was plausible that the IFR would rise in the situation of the NHS being overwhelmed, but felt it added assumptions to the modelling which were difficult to defend in detail.

209. In retrospect, spontaneous behaviour change and the dependence of IFR on healthcare demand were arguably both seen during the pandemic in the UK and other countries, though the details differed substantially from the scenarios modelled in Steven Riley's paper. IFR increased as healthcare systems became stressed (in the UK [NF/072 - INQ000262583] and elsewhere), but not to the degree that paper assumed. There were (mostly anecdotal) reports of significant spontaneous population behaviour change occurring around the peak of otherwise only partially mitigated pandemic waves in countries such as India, Brazil and Mexico, but this did not appear to lead to dramatically extended epidemics. Bulgaria is perhaps particularly worth noting as the country which suffered the highest level of excess deaths in the world in the first two years of the pandemic. Bulgaria saw little mortality before November 2020, since the country went into an lockdown in spring 2020 at a much earlier stage of their epidemic than did the UK. However, once NPIs were largely lifted by the autumn (and not significantly reimposed thereafter), the country saw an explosive epidemic wave in November and December 2020, with 0.25% of the population dying over those two months (measured by excess deaths) – more than the UK saw in the entire first 2 years of the pandemic. Bulgarian then saw a second, slightly slower wave of transmission caused by the more transmissible Alpha variant which led to the deaths of another 0.2% of the population over three months (March-May 2021), before seeing further epidemic waves caused by Delta and then Omicron. The timescale of these lethal epidemic waves (and those seen in several other countries) is not consistent with spontaneous behaviour change dramatically extending the duration of epidemics.

210. There was a more general email discussion between a large number of SPI-M-O members of the relative merits and feasibility of mitigation versus suppression on the 11th March [NF/134 - INQ000149011; NF/135 - INQ000149012; NF/136 - INQ000149013; NF/137 - INQ000149015; NF/138 - INQ000149016; NF/139 - INQ000149017; NF/140 - INQ000149018; NF/141 - INQ000149019; NF/142 - INQ000149020; NF/143 - INQ000149021; NF/144 - INQ000149022; NF/145 - INQ000149023; NF/146 - INQ000149024; NF/147 - INQ000149025; NF/148 - INQ000149026; NF/149 - INQ000149027; NF/150 - INQ000149028; NF/151 - INQ000149029; NF/152 - INQ000149030; NF/153- INQ000149031; NF/154 - INQ000149032; NF/155 - INQ000149033; NF/156 - INQ000149034; NF/157 - INQ000149035; NF/158 - INQ000149036; NF/159 - INQ000149037; NF/160 - INQ000149038; NF/161 - INQ000149039; NF/162 - INQ000149040; NF/163 - INQ000149041; NF/164 - INQ000149042; NF/165 -

INQ000149043], partly stimulated by Steven's paper. On 12th March, Jeremy Farrar, John Edmunds and I discussed by email where we thought government planning had got to at that point [NF/166 - INQ000149044; NF/167 - INQ000149045; NF/168 - INQ000149046; NF/169 - INQ000149047; NF/170 - INQ000149048; NF/171 - INQ000149049; NF/172 - INQ000149050; NF/173 - INQ000149051; NF/174 - INQ000149052; NF/175 - INQ000149053; NF/176 - INQ000149054; NF/177 - INQ000149055; NF/178 - INQ000149056; NF/179 - INQ000149057; NF/180 - INQ000149058; NF/181 - INQ000149059; NF/182 - INQ000149060; NF/183 - INQ000149061; NF/184 - INQ000149062; NF/185 - INQ000149063; NF/186 - INQ000149064]. At that time, I had started more systematic modelling of suppression options, focussing on examining the sensitivity of policy impacts to healthcare demand related thresholds for introducing and lifting intensive NPIs. We were also finalising clinical severity estimates (notably of IHR and ICU demand by age) that week [NF/048 - INQ000236303]. These two pieces of work were combined in the ICCRT Report 9 [NF/001 - INQ000049647], completed on March 15th.

211. The 13th March was pivotal in determining policy over the next 10 days. While the minutes give little sign of this, the SAGE meeting that day [NF/079 - INQ000109142] was sometimes tense and heated. In part this was because the first data from systematic NHS hospital surveillance had started to become available (though with significant data issues remaining – see section H), in part because unease came to a head about the logistical, political and ethical feasibility of a mitigation strategy which would overwhelm NHS capacity.

212. At that meeting, I very deliberately asked the NHSE representatives present whether there was any way the NHS could cope with the numbers of hospitalisations being envisaged under any of the mitigation scenarios previously reviewed by SAGE, even allowing for surging general bed and ICU capacity. This question got a very clear "No" in response, which was the first time NHSE had stated this at any SAGE meeting. Prior to that point, while NHSE staff had previously informally indicated this, it had not been stated "on the record" in my hearing. John Edmunds attempted to table a paper on the use of episodic lockdowns he brought with him to the meeting; an updated version [NF/187 - INQ000262628] was formally tabled at the 16th March meeting [NF/188 - INQ000075664]. The minutes of that meeting noted that "There is a risk that current proposed measures (individual and household isolation and social distancing) will not reduce demand enough: they may need to be coupled with more intensive actions to enable the NHS to cope, whether regionally or nationally". In terms of immediate

recommendations for intensification of NPIs, that SAGE meeting suggested that household isolation (i.e. all members of the household of a suspected case isolating for 14 days) be introduced as soon as possible.

213. Following that meeting, on 14th March, we were asked to prepare a slide deck for Patrick Vallance which illustrated the potential healthcare demand resulting from mitigation versus suppression policy options in a simple manner. I emailed a slide deck to GoS later that day and I believe this was shown to a senior ministerial meeting the following morning (15th March). The slide deck was included in the documentation for the 16th March SAGE meeting [NF/189 - INQ000119697]. Note that the SAGE website is incorrect in also attributing this slide deck to Cambridge University. Also, on March 15th, I had an email discussion with Chris Whitty and Patrick Valance [NF/384 - INQ000048089] about the level of NPIs which would be needed to avoid exceeding NHS ICU surge capacity. This made use of some of the results which would later be included in Report 9. I stated that "These conclusions are robust to uncertainty in severity and R_0", and I stand by that for the reasons stated in para 219 below.

214. As a member of NERVTAG as well as a SAGE participant, I also participated in discussions of the appropriate duration of case isolation at various times during the pandemic, including in an email discussion with Chris Whitty, Jonathan Van Tam and John Edmunds on 9th March 2020 [NF/190 - INQ000149065]. Chris Whitty asked whether I thought there would be a significant difference in the effectiveness of 7 versus 14 days of symptomatic case isolation – i.e. self-isolation of cases once they develop COVID-like symptoms. I felt there would be a small difference in the effectiveness of the two options. I based this view on the fact that most people infected with the original Wuhan strain only developed recognisable symptoms 4-6 days after infection, but that the generation interval (average time from a person being infected and them infecting others) was around 6-7 days. Thus isolation of 7 days would mean the average case would be isolating until approximately day 12 of infection, when infectiousness was much (likely 100-1000 fold) lower. While isolating for 14 days would clearly reduce transmission still further, I felt the gains would be small compared with the practical difficulties, cost and therefore perhaps lower adherence associated with longer isolation periods. In retrospect, I believe epidemiological data (contact tracing plus serial intervals observed in household studies) and virological data (on viral shedding over time, including in human challenge studies) have supported the views I expressed then. However, in SAGE and NERVTAG discussions at around the same time [NF/124 - INQ000109125], I advocated for 14

days of isolation of household contacts of suspected cases, since the contacts, if infected, would likely be at an earlier stage of their infection, given they had not developed symptoms.

March 16th–March 31st 2020 and Report 9

215. The week of 9th March also saw increasing public pressure, as indicated in the minutes of the 13th March SAGE meeting [NF/079 - INQ000109142], for the modelling of NPI policy options to be published. We therefore intensively worked on writing what would become Report 9 from 12-16 March. It was publicly released at 5pm on 16th March [NF/001 - INQ000049647] but shared with government and presented to SAGE [NF/191 - INQ000197149] earlier that day.

216. While Report 9 summarised much of the earlier SAGE modelling of mitigation, its aim was primarily pedagogical: to set out the fundamental difference between mitigation and suppression strategies in terms of their effects (on transmission, health impact and policy duration), and to highlight that any strategy which failed to bring R below 1 would likely result in healthcare demand exceeding healthcare capacity many-fold over. I recognise that our conclusion that "epidemic suppression is the only viable strategy at the current time" clearly proposes a policy strategy. However, in a UK context, I felt it reflected a decision which had already fundamentally been made the previous day in the UK, and rather earlier by many other European countries. We also noted that "the social and economic effects of the measures which are needed to achieve this policy goal will be profound" and that "it is not at all certain that suppression will succeed long term".

217. In preparing Report 9, we chose to use the individual-based model (CovidSim) because that was the model we had used for all prior modelling of NPI policy options for SAGE. However, we could have equally used a simpler compartmental model to illustrate the fundamental differences between suppression and mitigation policies and the healthcare demand and potential mortality under each; CovidSim had the advantage of being able to mechanistically model NPIs such as household quarantine and school or workplace closure (in that the model explicitly incorporates those social units), but the disadvantage of higher parametric and computational complexity. In later work, such as Report 12 (which modelled mitigation and suppression strategies for every country [NF/002 - INQ000262593]), we switched to a simpler compartmental model and analytical calculations (of epidemic "final size"), and obtained almost identical results for the modelling of the UK as reported in Report

9. In addition, the LSHTM modelling, again using a compartmental model, also obtained very similar results [NF/192 - INQ000262582; NF/187 - INQ000262628].

218. "Lockdown" was not modelled in Report 9, and the word is only used once. We also did not discuss the issue of voluntary versus mandatory measures, though the closure of schools and universities is clearly in the latter category. The most intensive policy modelled involved suspect cases and their household members isolating, school and university closure, a 75% reduction in social contacts outside the household or workplace, and a 25% reduction in workplace contacts. This set of interventions was more akin to what the Prime Minister announced over the week of 16-20 March than the more intensive lockdown announced on 23rd March.

219. I stand by the fundamental conclusions of Report 9, and the results presented. There were limitations to the modelling included, but those either reflected the time pressures we were working under, or the paucity of some data used to inform model parameter estimates. The key limitations were:

a. Children: while the model included children being much less likely to develop severe disease and being somewhat less likely to develop symptoms, in the absence of any data on infection attack rates by age at that point, it made the conservative assumption that children were as susceptible to infection as adults. In retrospect, this was not the case for children under 14 for the variant of the virus which caused the first UK wave of infection, though later variants (from Alpha onwards) infected younger children more efficiently. The assumption around children meant that the model results likely over-estimated the impact of school closure (particularly primary school closure) on transmission. I note that the role of children in transmission was the topic of much SAGE activity and discussion for the remainder of 2020.

b. Shielding of the elderly: for the mitigation scenario presented, we assumed that contact rates in the over 70s could be reduced by 75% while the rest of society continued as normal. I felt this was an optimistic assumption (see para 193), but retained it largely to show that even under an optimistic assumption around the effectiveness of shielding, projected hospital demand under the mitigation scenario would still exceed hospital capacity many-fold over. We did not include shielding in the modelling of suppression

options, but rather modelled a blanket 75% reduction in social contacts for the entire population.

c. Lack of detailed transmission data for COVID-19: detailed and reliable data on the transmission rates of COVID-19 in different settings (notably households) would not become available for weeks to months after Report 9 was prepared. Thus we assumed the relative risks of transmission in different settings would be comparable to influenza and other respiratory viruses (and also consistent with what is measured in social contact surveys such as COMIX), albeit scaled by the overall value of R_0 assumed for COVID-19. Data collected later largely supported this assumption, with the exception of assumed rates of transmission in primary schools (see point *a* above).

d. Limited uncertainty/sensitivity analysis: the computationally intensive nature of the CovidSim model meant we had limited ability to undertake extensive sensitivity analysis in the time available, and so focussed on exploring how policy impacts varied with R_0 and with the healthcare demand related triggers used to start and stop NPIs. Had we had more time, we would have explored a larger range of scenarios for the assumed effectiveness of different NPIs, most notably social distancing and household quarantine. I would note that the projected numbers of hospitalisations and deaths in Report 9 scale linearly with the assumed overall infection hospitalisation ratio (IHR) and IFR, respectively; hence projected mortality for an IFR of half that assumed would just be half of the numbers published. Therefore the conclusion that a mitigation strategy would overwhelm NHS capacity was valid even if IFR had been 0.3% or even lower. We therefore did not view exploration of sensitivity to severity parameters as essential, though we could have made the point about linear scaling explicit in the text (I believe it was clear to SAGE and SPI-M-O participants).

e. Epidemic growth rate: The range of R_0 values explored in Report 9 (between 2 and 2.6) reproduced epidemic doubling times (in the absence of any control measures) between 3.8 days (R_0=2.6) and 5 days (R_0=2). The epidemic doubling time seen in the UK in March 2020 turned out to be at the low end of this range, as estimated from the first reliable surveillance data available (the NHS Sitrep – see section H). In retrospect, one cause of the higher-than-expected growth rate was viral evolution – the clade of virus which dominated the European first wave had a D614G mutation in the Spike protein not present in the Wuhan virus (see section K). Other factors may have also contributed. Thus, our later estimates of R_0 using NHS Sitrep data for the week 16-23 March (the earliest reliable data

available) were closer to 2.8 to 3, depending on what was assumed about the effect of control measures introduced prior to 23 March. This discrepancy had only a small effect on mortality and hospital demand projections for the unmitigated and mitigated scenarios, but a somewhat larger effect on suppression scenarios – in particular meaning that only the most intensive NPI policy modelled in Report 9 could be expected to achieve suppression (*i.e.* reduce R to below 1).

f. Stage of epidemic: Given the limitations of UK (and US) surveillance at the time the work was conducted, Report 9 presented scenario modelling rather than precise predictions of epidemic trajectory. For Figures 2 and 3 in the report, the horizontal axis was calibrated such that there were 20 deaths in GB (Northern Ireland was not included) by 14th March, matching reported numbers up to that date, but we were aware that surveillance was at best partial at that time, meaning a substantial number of deaths might have been missed. In reality, the data available from 16th March onwards indicated that the UK epidemic had progressed substantially further than we had anticipated (see section H). This did not affect projections of total deaths for the unmitigated and mitigated scenarios but it did mean that the highest ICU occupancy triggers for initiating NPIs we modelled in the suppression scenarios (400 cases entering ICUs in the previous 7 days) had already been exceeded by March 22nd 2020.

g. Hospitalisation parameters: At the time of writing the report, we had very little UK-specific data on the IHR, the proportion of hospitalised patients who would require ICU, or lengths of hospital stay for COVID-19 patients. As discussed above, on 14th March our best estimate of the IHR was 4.4%, with 30% of patients requiring ICU – the latter value only being adopted that day after discussions with NHSE colleagues. In reality, IHR in the UK in the first wave was closer to 3%. In addition, ICU bed occupancy was over-estimated in Report 9, as the actual median duration of stay in ICU in the UK first wave turned out to be substantially shorter (especially for patients who later died) than the 10 days we assumed for Report 9 (based on expert clinical judgement). In part this discrepancy may reflect demand-related triaging decisions made in NHS hospitals in March and April 2020. Also, in the first wave in the UK, a substantial proportion of COVID-19 deaths occurred outside hospitals – over 3000 in private homes (with COVID-19 recorded on the death certificate), and many more in care homes. While we had assumed in Report 9 that all severely ill COVID-19 cases would be hospitalised. Thus while the suppression scenarios in Report 9

underestimated first wave mortality (due to suppression in the UK starting later than any of the scenarios modelled), they over-estimated peak ICU (and, to a lesser degree, general hospital ward) occupancy. The degree of discrepancy made no difference to the conclusion that mitigation strategies would overwhelm NHS capacity many-fold over.

h. Population behaviour: we did not attempt to include any spontaneous change in population behaviour in response to the epidemic, since validated quantitative models of such behavioural responses are lacking. See also para 46, 49 and 510.

220. The inclusion of an unmitigated scenario in Report 9 has been criticised by some as unnecessary and/or alarmist. However, it is standard practice in epidemiological modelling of disease intervention strategies to include a counterfactual of what would happen in the absence of intervention, as this is the only way to quantify the potential benefits of acting. In reality, our unmitigated scenario was likely optimistic in its projection of mortality. Deaths would likely have exceeded the number quoted if the NHS had really been allowed to be completely overwhelmed, given IFR estimates implicitly assumed availability of hospital treatment.

221. Following discussions within government on the 15th, the 16th March SAGE meeting [NF/188 - INQ000075664] also saw a significant pivot towards advising rapid introduction of more intensive NPIs. This was driven by two factors: (a) acceptance that mitigation strategies would lead to overwhelming healthcare demand and that such an outcome was now viewed by government as unacceptable, and (b) preliminary surveillance data from ICUs then available (see section H) which suggested the epidemic was significantly more advanced in the UK than previously thought. School closure was mentioned as potentially being necessary, though with significant and justified caveats.

222. On 17th March I developed relatively mild (dry cough and a slight fever for two days) COVID-19 symptoms. My infection was later verified through PCR testing, ordered in light of the meetings involving key officials I'd attended in the previous week. I self-isolated for two weeks from 17th March, but continued to work. However, this meant I had to dial into the following two largely in-person SAGE meetings (18th and 23rd March), limiting my ability to hear discussions or participate (the AV arrangements at that time being fairly awful). In addition, I was feeling quite unwell on the 18th, which further limited my participation.

223. My limited recollection of the SAGE meeting of 18th March [NF/193 - INQ000075778] was that it focussed on discussing what surveillance data could tell us about the stage of the epidemic in the UK (including regional variation) and what additional NPI measures might be needed above those announced on 16th March to avoid NHS capacity limits being breached. ICCRT contributed one of three modelling papers on the impact of school closures [NF/194 - INQ000231032] to that meeting. Consensus papers on the topic were also submitted by SPI-M-O [NF/195 - INQ000074903] and SPI-B [NF/196 - INQ000262621], the former of which highlighted the many uncertainties, including about the contribution of children to transmission. The meeting also discussed whether more intensive measures were needed sooner in areas of the country (notably London) where the epidemic was more advanced.

224. On 18th March, the government announced that all English schools, colleges and early-years units would close from 23rd March, other than for children of key workers [NF/197 - INQ000052716]. In doing so, the government followed the actions of nearly every other European country (though Sweden only closed secondary schools and the Netherlands initially kept schools open for those taking examinations). Even now, it is hard to assess what the effect of school closure was on transmission rates in the first UK COVID-19 wave; schools were kept open in the second (November 2020) England lockdown and that lockdown succeeded in reducing R to below 1, but not as low as the 0.6-0.7 value reached in the first lockdown. However, by November 2020, there had been substantial investment in making schools and other public facilities "COVID-safe". Conversely, the lower apparent effectiveness of the second lockdown may have also been influenced by the growth of the more transmissible Alpha variant at that time. I note that any appreciable reduction in effectiveness of the first lockdown would have increased the total number of hospitalisations and deaths seen in the first wave [NF/022 - INQ000262573]. Overall, my retrospective judgement is that the closure of primary schools in the first wave probably had a limited impact on transmission overall, but closure of secondary schools and universities likely had an impact comparable to what was reviewed at the 18th March 2020 SAGE meeting. However, given the lack of good data on transmission in children (and asymptomatic transmission more generally) at the time the decision had to be made, I believe primary school closure was justified from a precautionary perspective.

225. Much of the SPI-M-O discussion during the week of 16th March focussed on better characterising the stage of the epidemic and its growth rate (*i.e.* doubling time or R). I discuss this topic in section H.

226. The SAGE meeting on 23rd March [NF/198 - INQ000052717] concluded the epidemic growth rate was faster than previously anticipated, with the 20th March SPI-M-O consensus statement [NF/199 - INQ000228591] giving doubling times in the range 3-5 days. My own assessment was that the true doubling time was approximately 4 days, but others argued it might be as short as 3 days. I discuss this complex issue in section H. SAGE also discussed the extent to which measures introduced prior to that date had succeeded in reducing population rates sufficiently to bring R to below 1 in the timescale needed to prevent NHS capacity limits being exceeded, especially in London. There was a general consensus that additional measures were needed, especially in London. There was no new work from ICCRT presented at that meeting.

227. I conclude this section with some brief reflections on uncertainty in the epidemiological modelling informing SAGE deliberations in February and March 2020. I have already discussed model uncertainty more generally in section D. While all the NPI modelling reports considered by SAGE in that time period were concise, I believe that Chris Whitty and Patrick Vallance had a good appreciation of the uncertainties in our (and other SPI-M-O groups') modelling of NPIs from the caveats in the documents themselves, discussions within and outside SAGE, and from the SPI-M-O consensus statements. I also tried to convey the uncertainties in discussions with other civil servants, as I know Patrick Vallance and Chris Whitty did. My view was, and remains, that in a rapidly-moving crisis where time is at a premium, "best efforts" analysis and modelling – particularly by multiple groups at once, to give a sense check – is far preferable to no analysis and modelling at all. I also note that the considerations and conclusions of SPI-B on behavioural responses to the pandemic and to NPIs were equally uncertain – almost inevitably, given the unprecedented nature of the crisis.

H. Tracking the pandemic: COVID-19 surveillance in the UK

Access to UK data by SPI-M-O groups

228. The table on the following page summarises the main sources of surveillance data in England during the pandemic. Similar systems were established in the devolved administrations, aggregated data streams from which SPI-M-O groups also gained access.

229. ICCRT access to most data sources was provided under a data sharing contract between PHE and MRC GIDA enacted on 5th March 2020. This contract was facilitated by the SPI-M-O secretariat and distributed at the same time to all SPI-M-O groups. Separate data sharing agreements were later signed for access to CO-CIN and ONS data (both facilitated by SPI-M-O), and for COG-UK data.

Surveillance up to 23rd March 2020

230. COVID-19 surveillance in the UK went from being far from adequate prior to 16th March 2020 to being one of the most (if not the most) comprehensive systems put in place by any country.

Table: UK COVID-19 surveillance datasets used by ICCRT for SAGE/SPI-M-O work

Start date column shows date ICCRT first had access.

Data type	Dataset	Details	Shared by	Start date
Cases	FF100 linelist and COVID-19 main linelist	Anonymised individual-level data on all people testing positive for COVID infection, giving an anonymised identifier to allow linkage to other datasets. Both Pillar 1 and Pillar 2 positive tests were included. From January 2021 episode number also included. Until 18/3/2020 included a list of deaths - later moved to a separate dataset. FF100 largely discontinued after March 2020.	UKHSA (formerly PHE)	06/03/20
	RCGP swabbing	GP sentinel surveillance data - discontinued after 26/06/2020	UKHSA	16/03/20
	Negative test linelist	Anonymous linelist of all negative tests, by LTLA and date of test	UKHSA	01/10/20
	Reinfections linelist	From Nov 2021 - Jan 2022, a separate anonymous linelist of cases known to be a second case, linked to the first case. Superseded when UKHSA data files moved to an episode-based approach	UKHSA	04/11/21

Hospital demand	NHS COVID-19 sitrep	Daily aggregated returns from all NHS trusts giving Trust level counts of numbers of patients admitted with COVID-19, numbers currently occupying general ward beds, ICU beds or ventilator beds. Data provided at Trust level.	NHSE	16/03/20
	MHLDA sitrep	Daily aggregated returns similar to main NHS Sitrep specifically from mental health and learning-disability health settings	NHSE	15/05/20
	COVID-19 Hospitalisation in England Surveillance System (CHESS)	Anonymous individual level hospital patient data providing length of stay and other clinical data. Unlike NHS Sitrep, provided individual-level patient data, but coverage was less complete.	UKHSA	17/03/20
	ISARIC4C COVID-19 Clinical Information Network (CO-CIN)	UK component of the UK-led but international ISARIC network. Research study spanning many UK hospitals, individual level hospital patient data providing length of stay and other clinical data. More detailed than CHESS, though covered fewer hospitals.	ISARIC	26/03/20
	SUS_ECDS	Dataset of individual COVID-related hospital episodes, identified by the same anonymous identifier as the main line list.	UKHSA	13/09/21
Deaths	COVID-19 death line-list	Individual level dataset of COVID-associated deaths using the same anonymous identifier as for the main line-list.	UKHSA	18/03/20
	ONS deaths	Deaths due to COVID-19 by age group, local authority and date of death	ONS	12/05/20
Antibody prevalence	PHE serology	Antibody prevalence surveys of blood donors and other groups.	UKHSA	21/04/20
	REACT-2 study	Large-scale random population antibody prevalence survey	Imperial College	13/08/20
Infection prevalence	ONS infection survey	Population survey of SARS-CoV-2 infection prevalence (PCR positivity)	ONS	14/05/20
	REACT-1 study	Population survey of SARS-CoV-2 infection prevalence (PCR positivity)	Imperial College	10/07/20
Variants	COVID-19 Genomics UK Consortium	SARS-CoV-2 viral genome sequences with metadata (e.g. age, LTLA, date of sample). Later supplemented with lineage assignment.	COG-UK	23/03/20
	Variant linelist	Subset of main linelist giving the genetic variant of the virus identified through genotyping or genetic sequencing.	UKHSA	29/04/21
	SGTF (S-gene target failure) linelist	Subset of main linelist of tests processed using the TaqPath PCR test used at the Lighthouse labs, giving the PCR CT values for each genetic target of the PCR test used to diagnose a case.	UKHSA	30/12/20
	Contact-tracing data	Extract from the Test and Trace database providing data on cases linked by contact tracing and data on the variant causing infection	UKHSA	29/3/21
Vaccination	Immunisation line-list (NIMS)	Linelist of every COVID-19 vaccine dose given in England using the same anonymous identifier as the main line-list.	UKHSA	22/01/21

231. Prior to 25th February 2020 (and in functional terms, 16th March), it was moot whether the UK really had a surveillance system for COVID-19, if one defines surveillance as a systematic approach to tracking the level of infection present in the country. Rather, the UK had a protocol for testing people matching a certain clinical case definition who had a recent history of travel to a designated high-risk region or recent contact with a laboratory confirmed cases.

232. This early focus on testing of travellers was an essential component of the Contain phase of the UK government response. As I discussed previously in section G, I felt the Contain phase never had any significant chance of preventing infection entering the country or even significantly slowing its establishment here. However, the effectiveness of the policy was further impaired by the limited testing capacity available, this being a constraint on how risk averse we could be in designating high-risk regions meriting testing of travellers. At the NERVTAG meeting of 28th January 2020, I questioned the aim behind the then approach to which overseas areas were classified as high-risk. The minutes state that "CMO and DCMO commented that there would need to be a pragmatic balance between the sensitivity of the case definition and the capacity of the health system under the winter pressures." [NF/082 - INQ000047820].

233. The minutes of that NERVTAG meeting also record, in relation to the epidemic in China, that I "noted that the case numbers seem to be doubling every three to four days. There is no evidence of it slowing so far but it is hard to tell given the delay in the incubation period and how long it is taking to confirm cases at the moment".

234. As requested at that NERVTAG meeting, I emailed an assessment of the risk of importing infections from different Chinese provinces to Jonathan Van Tam and Chris Whitty later on 28th Jan 2020 [NF/200 - INQ000149066; NF/201 - INQ000149067; NF/202 - INQ000149068; NF/203 - INQ000149069; NF/204 - INQ000149070; NF/205 - INQ000149071; NF/206 - INQ000149072; NF/207 - INQ000149073; NF/208 - INQ000149074]. This suggested that due to the higher passenger numbers, we would expect more infections to be imported from Beijing and Shanghai than Hubei, despite the much higher case numbers in Hubei.

235. Testing of travellers was extended from symptomatic passengers from Wuhan to passengers from anywhere in China (plus a number of other countries) on 7th February 2020 [NF/209 - INQ000051974]. The list of overseas areas designated as high risk was frequently

expanded from that point on, meaning it was difficult to interpret what trends in numbers of diagnosed cases really implied for the underlying growth rate of the global pandemic. Generally, when testing is expanded over time, case numbers will be expected to rise even if the incidence of infection is static. In the context of January-March 2020, this meant there was a risk of over-estimating the growth rate of the epidemic from naïve analyses of the total case numbers reported by different countries.

236. Given this potential bias, and that I was certain from early February 2020 that infection would become established in the UK (if it wasn't already), I urged at SAGE meetings that systematic surveillance be established within the UK as soon as possible. The minutes to the 11th February meeting of SAGE include a statement that "The UK has 8 confirmed cases, all of whom acquired the virus overseas" [NF/210 - INQ000075784]; I felt that since this statement was axiomatically true (the criteria for testing requiring recent travel from high-risk area), it was uninformative and risked complacency. In my email exchange with Chris Whitty on 21st February I stated *"So I am reminded of the BSE enquiry conclusion of >30 years ago that governments should not view absence of evidence as evidence of absence. I think the epidemiological data - while patchy - suggests that we should be acting on the assumption that sustained transmission is happening now in the UK. And only de-escalate if we can demonstrate - through surveillance - that it is not"* NF/093 - INQ000148969. I think subsequent events vindicated my view.

237. Surveillance limitations are referred to in the minutes to the 20th February SAGE meeting [NF/094 - INQ000262616] and I submitted papers analysing the potential sensitivity of different sentinel surveillance designs to the 25th February SAGE meeting [NF/096 - INQ000087503]; NF/094 - INQ000262616][124].

238. While comprehensive testing clearly gives the most information, when resources (such as testing capacity) are limited, well-designed sentinel surveillance is the preferred approach to derive unbiased estimates of the incidence (new cases per day per capita) of an infection in a population. Sentinel surveillance involves the recruitment of a random network of healthcare providers (GPs, hospitals or ICUs) to systematically collect samples (for later testing) from patients meeting predefined clinical criteria. From the data collected from the patient population covered by the network, one can extrapolate trends to the rest of the population. Prior to the pandemic, PHE and the Royal College of General Practitioners (RCGP) Research and Surveillance Centre in Birmingham ran a GP-based sentinel surveillance network for a

number of respiratory infections, notably influenza, the data from which formed an important part of the country's monitoring of those infections. A hospital-based surveillance system – the UK Severe Influenza Surveillance System (UISSS) also operated prior to the pandemic.

239. On 25th February 2020, PHE announced a new ICU-based sentinel surveillance system [NF/211 - INQ000086641], involving 8 hospitals at that stage (*i.e.* 10% of ICUs nationally), all of which were part of UKISSS. SAGE participants were also informed at the meeting on 25th February that sentinel GP surveillance (via the RCGP network) came on stream on the same date. We were informed that it was expected then that these two systems would likely result in about 300 samples being tested in the first week. I do not know exactly how many GP practices had been recruited to the GP network.

240. I became aware (from informal conversations with Jonathan Van Tam) that one COVID-19 case had been detected through the RCGP network by 29th February, and that 3 RCGP cases and one ICU case were confirmed on 4th March. All these cases were unlinked to travel. To my recollection, this was the first definitive evidence indicating that community transmission was ongoing in the UK.

241. We (together with the LSHTM, Lancaster, Cambridge and Manchester SPI-M-O groups) received the first detailed data on UK COVID-19 cases from PHE on 6th March 2020. Data updates were received mostly daily thereafter, initially through an encrypted email service and later through a secure server. These data were in linelist format – a spreadsheet with a row for each case. Initially we received what was called the "first few hundred" (FF100) dataset, which included relatively detailed (though often incomplete) data on each case, including some information on contact tracing. The intention of the FF100 dataset (the creation of which was part of UK pandemic preparedness planning before 2020) was less to track the epidemic, than to provide detailed information on a subset of cases to allow estimation of epidemiological parameters (e.g. mean time from symptom onset to hospitalisation).

242. Perhaps for that reason, the 6th March FF100 dataset was incomplete, as it listed just 116 cumulative cases compared with the total of 164 cases which had been reported in the UK by that date. Of those 116, all but 8 were in travellers or contacts of travellers. Of the remaining 8 cases, 6 had been detected through sentinel surveillance (3 in hospital) and the other two were contacts of one of the latter 6 (detected via contact tracing).

243. On 10th March, PHE launched a new public COVID-19 dashboard [NF/212 - INQ000262578] which included the geographic distribution of COVID-19 cases and deaths, updated daily. Also from that date, a second line-list dataset was shared with SPI-M-O groups, called the "COVID-19 Anonymised linelist", which I think (but don't know) was derived from the PHE dashboard inputs. The anonymised linelist included nearly all of the publicly announced cases, but contained less data on each case than the FF100 dataset (which, for the avoidance of doubt, was also anonymised). In particular, the FF100 dataset provided substantially more information on hospitalisations and whether a case was travel-linked or sporadic (*i.e.* not linked). We used a combination of both datasets through to mid-March but relied increasingly more on the full anonymous linelist as it became more complete. The FF100 dataset was never expanded beyond 416 cases, all of which were detected before the end of March 2020.

244. The incompleteness of data in the anonymous line list up to 16th March severely limited its usefulness for analysis and modelling. In particular, data on travel history and hospitalisation status were missing for most cases. For example, of the 745 cases included in the 13th March line-list, only 224 had (international) travel history status completed, and there was no information on hospitalisation status for any of the cases reported from 9-13 March. An additional challenge was identifying which cases had been detected via the different sentinel surveillance networks, which through testing of travellers, and which through contact tracing. This distinction was critical, as only the sentinel surveillance gave a relatively unbiased picture of transmission within the country.

245. Conversely, there was an approximately 3-day delay in adding recent cases to the more complete FF100 line list at that time.

246. We only gained access to a separate data file for the GP swabbing system from 16th March 2020, though GP participation in that system apparently dropped rapidly due to the deep clean protocols that practices had to follow and the contact tracing which occurred every time a positive case was found (prior to 16th March, at least).

247. My view then and now is that resourcing of data capture, linkage, analysis and distribution was woefully inadequate in PHE between January and March 2020. The responsibility for organising case data for sharing with SPI-M-O groups in mid-March seemed to fall on a handful of people in PHE, all of whom were working 18+h days, 7 days a week by 13th March. Indeed, I was so concerned that I sent ICCRT staff to see if they could help the PHE team at Colindale

on 12th March. After they returned, they described such a chaotic situation in terms of data organisation that I emailed Sharon Peacock and Nick Phin (the PHE staff in charge of COVID-19 surveillance) and Jonathan Van Tam with my concerns that evening and forwarded the same email to Patrick Vallance later [NF/213 - INQ000148802; NF/214 - INQ000148803; NF/215 - INQ000148804; NF/216 - INQ000148805; NF/217 - INQ000148806].

248. The statement in the March 13th SAGE meeting minutes that "Owing to a 5-7 day lag in data provision for modelling, SAGE now believes there are more cases in the UK than SAGE previously expected at this point" [NF/079 - INQ000109142] reflects the delays in updating the FF100 dataset (which gave hospitalisation information) referred to above, plus my (together with Graham Medley's and John Edmunds') increasing realisation about how incomplete (in terms of information on hospitalisation and travel status) the data shared with SPI-M-O groups then was – as illustrated by the 5 fold increase in cases that were known to be unlinked to travellers and had been hospitalised between 6th and 11th March.

249. More generally, surveillance expanded rapidly that week (beginning 9th March 2020). Of particular note was the expansion of hospital surveillance and consequent refocussing of testing capacity from largely testing travellers and their contacts on 9th March to prioritising the testing of hospital patients by the 16th of March. However, at the time, I and my team had only a partial view of the initiatives underway. Testing capacity was also being scaled up substantially at this time, though again we did not have detailed data on this at the time. In the week 2-8 March 2020, an average of 1956 COVID-19 tests per day were performed, rising to an average of 5255 per day in the week 9-15 March.

Hospital-based surveillance

250. On 11th March, the COVID-19 Hospitalisation in England Surveillance System (CHESS) was established, with a letter sent by PHE to all NHS trusts requiring mandatory daily reporting of all hospitalised cases (aggregated totals) and individual data on all cases in ICUs from the following day [NF/218 - INQ000119666]. I was informed of the details of this system by Andrew Jackson at NHSE on 13th March [NF/219 - INQ000148819; NF/220 - INQ000148821; NF/221 - INQ000148823; NF/222 - INQ000148825]. We and other SPI-M-O groups first received data from CHESS on 16th March 2020.

251. Sometime that week (i.e. starting 9th March), NHS England also instituted what I think was a separate system for monitoring COVID-19 related healthcare demand. This was intended

for demand management and planning purposes rather than systematic surveillance, but provided daily total numbers (by NHS Trust) of newly admitted or diagnosed patients, bed occupancy data and ventilator use. This system is termed the NHS COVID-19 Sitrep in the table above. I first became aware of the system's existence on the morning of 16th March, I think from a discussion with Jonathan Van Tam, though I can't be sure. Jonathan Van Tam authorised the data from that system to be shared with SPI-M-O groups, and I then emailed Andrew Jackson at NHSE (copying Jonathan) requesting access [NF/223 - INQ000148809; NF/224 - INQ000148810; NF/225 - INQ000148811; NF/226 - INQ000148812; NF/227 - INQ000148813; NF/228 - INQ000148814; NF/229 - INQ000148815; NF/230 - INQ000148816; NF/231 - INQ000148817; NF/232 - INQ000148818]. Andrew Jackson sent me the first data from that system later that day, which provided NHS Trust level on the number of newly diagnosed COVID-19 patients (plus general ward and ICU bed occupancy) each day between 9-15 March (though data for the 15th were very incomplete). I don't know whether these data had been collected each day during the previous week or whether some had been collected retrospectively. The system was substantially revised over the weekend of 14-15th March, with more detailed data becoming available as of 17th March.

252. MRC GIDA received the NHS COVID Sitrep every day from 16th March. I am aware that not all SPI-M-O groups obtained access from the same date as we did (and I think LSHTM, but this would need to be checked with John Edmunds). My email to Andrew Jackson and Jonathan Van Tam referred to above made it clear that Jonathan had authorised access for SPI-M-O groups. I don't know why there was a delay in providing wider access, but there may have been confidentiality concerns which were mitigated in our case by the fact we had already been working very closely with the NHSE analytics team throughout March.

253. While by late March the CHESS system was providing invaluable detailed data on the progression of patients in hospitals (e.g. length of stay, proportion being admitted to ICUs), I heard anecdotal reports that some NHS trusts struggled with the data entry requirements, at least in the first few weeks it was operational – meaning the system did not provide as complete a picture of rising hospital demand as was intended.

254. Hence from 16th March on, the NHS COVID-19 Sitrep provided the best available compromise of timeliness and completeness in providing a summary picture of COVID-19 associated healthcare demand, and, until mass community testing started some months later,

of the progress of the epidemic in the country. That said, prior to 21st March, returns were missing from approximately 10% of NHS Trusts.

Interpreting epidemic trends 6-23 March

255. Apparent growth rates in case numbers need to be interpreted cautiously when surveillance is expanding rapidly, as was the case across Europe during March 2020. A number of public websites at that time (and one or two SPI-M-O groups) plotted cumulative case numbers on a log scale, and associated linear regression lines suggesting several countries (including the UK) were seeing a 3 day doubling time. For instance, the increase in the cumulative total reported number of UK cases from 319 on 9th March to 1543 on 16th March suggested a 3.1 day doubling time. My view is that these analyses somewhat overestimated the true underlying growth rate due to the rapid scaling up of surveillance during March 2020. In addition, there are significant statistical issues with estimating epidemic growth rates from cumulative case numbers, and in using the report date of a case to make such estimates (rather than the symptom onset date or date of test).

256. However, in the absence of any other data, I calibrated the simulations used in Report 9 [NF/001 - INQ000049647] so that the maximum R_0=2.6 unmitigated epidemic scenario explored in that report reproduced the 3.8-day epidemic doubling time we estimated from daily cases reported in the UK in the 14 day period 28th February to 12th March 2020, using the UK linelist we received from PHE on 12th March. A two-week period is the minimum needed to reliably estimate a doubling time, given day-of-week biases in case reporting. The R_0=2.4 scenario shown in Figures 2 and 3 of that report gave a doubling time of 4.2 days. The simulations assumed a slower (5 day) doubling time of imported infections into the country, to represent the cumulative effect of increasing border restrictions and decreasing international travel.

257. Analysing the daily counts of newly diagnosed COVID-19 cases in English hospitals (from the NHS COVID Sitrep) for the 7 days from 16th March to 22nd March 2020 gives a doubling time estimate of 3.7 days. Given that there was an average 4-day delay from symptom onset to hospital admission, and that the average incubation period was approximately 5 days, this doubling time applies to the underlying epidemic growth rate 9 days earlier (7-13 March).

258. However, since daily data are subject to day-of-week biases (people are more likely to present on some days rather than others, and the rapidity of testing also varies), estimates of

growth rate made from a single week's data can be unreliable. It is preferable to assess growth rate over a minimum of a 14-day period. Comparing the 2836 newly diagnosed patients reported in the NHS Sitrep in the week 16-22 March 2020 with the 9117 reported for the week 23-29 March gives a doubling time estimate of 4.2 days – though even allowing for the average 9-day delay from infection to admission, this latter estimate may be affected by slowing of the epidemic caused by the policies introduced in the week of 16th March (but likely not by the 23rd March lockdown).

259. The most important new information provided by the initiation of systematic hospital-based surveillance was that the epidemic in the UK had progressed substantially further than was previously thought, based on the earlier traveller-focussed surveillance. While there were only 8 (3 hospitalised) non-travel related cases in the 6th March 2020 FF100 linelist, this rose to 43 (16 in hospital) by the 11th March FF100 list (the anonymous linelist of the same date had less complete hospitalisation data and showed even fewer non-traveller hospitalised cases). The (still partial coverage) NHS COVID Sitrep of 13th March recorded 350 COVID-19 patients in hospital (though we only received the latter data on 16th March). The NHS Sitrep of 21st March (the first with almost complete returns) reported 2156 COVID-19 patients in hospitals.

260. This 700-fold increase in hospitalised cases in over just two weeks partly reflects the approximately 4 doublings (16-fold) increase in the size of the epidemic, but also indicates the degree to which hospitalised cases were under-ascertained (*i.e.* missed) prior to full establishment of the NHS COVID Sitrep system.

261. Had the 11th March FF100 linelist figure of 16 hospitalised cases been accurate, then we would have "only" expected 160 hospitalised cases by 21st March, even making the pessimistic assumption of a 3 day doubling time. Hence surveillance data available on 11th March was missing more than 90% of hospitalised cases, and likely an even higher fraction of mild cases. The level of under-ascertainment was even worse going back to 6th March. This discrepancy is equivalent to approximately 2 weeks of epidemic growth – *i.e.* in reality, the UK epidemic was approximately two weeks ahead of what was 11th March data indicated.

262. At some time before the March 10th SAGE meeting (I don't recall when), John Edmunds and I came up with a simple approximate approach for getting a sense of the scale of under-ascertainment of cases, later formalised by both the LSHTM group [NF/233 - INQ000262595]

and ICCRT [NF/018 - INQ000262556]. This involved using reported deaths or ICU cases as a proxy for the underlying number of infections, on the basis that very severe cases were more likely to be tested. Assuming a 1% IFR, every death corresponds to approximately 100 underlying infections. But given an approximately 2-week delay from symptom onset to death, that measure of infections corresponds to two weeks previously, over which time we assumed (conservatively) that epidemic had grown 10-fold in size. A similar argument could be used to scale ICU cases, albeit with greater uncertainty, due to the less precise estimates of the proportion of infected individuals who might require ICU admission. This very approximate calculation underlay the statement in the 10th March SAGE meeting minutes that "...the UK likely has thousands of cases – as many as 5,000 to 10,000" [NF/124 - INQ000109125; NF/019 - INQ000228162]. However, in retrospect that range was almost certainly a substantial underestimate, given (based on the arguments above), it is highly likely that many (if not most) COVID-related deaths were missed in the UK prior to 11th March 2020.

263. While we always expected that the majority of cases were being missed with the surveillance in place prior to March 9th, SAGE participants (including me) only gradually became aware of the full extent of this discrepancy as the data accumulated. These concerns are mentioned in the 13th and 16th March meeting minutes [NF/079 - INQ000109142; NF/188 - INQ000075664],[151] and I believe were key drivers of the acceleration of policy decision-making from 13th March onwards.

264. As we expected, there was substantial regional variation in the progress of the epidemic by mid-March 2020 – with London being the region which was significantly ahead (given its greater global connectivity). I tracked this from 18th March using NHS Sitrep data and emailed Patrick Vallance an assessment of the implications for NHS demand in different regions on 20th March [NF/219 - INQ000148821; NF/220 - INQ000148822; NF/221 - INQ000148823], an updated version of which was considered (with similar submissions from other groups) by SPI-M-O on 23rd March. My 20th March analysis indicated that the NHS in London would exceed surge capacity within 2 weeks at the then rate of growth in admissions. The SPI-M-O consensus statement of 20th March [NF/199 - INQ000228591] agreed with this conclusion, and it was reflected in SAGE conclusions on 23rd March 2020 [NF/198 - INQ000052717]. I believe these conclusions were pivotal in the announcement of the first lockdown on 23rd March, though I was not party to policy-making discussions within government.

265. The RWCS and NHS planning assumptions agreed by SAGE on 6th and 10th March 2020 were not affected by the surveillance limitations detailed above, as neither attempted to say anything about the date when the epidemic might peak or the stage of the epidemic the UK had reached at the time they were agreed. At the time they were being drawn up, the official UK government policy was still containment, and official documents referred to "if" rather than "when" local transmission became established (e.g. the March 3rd COVID-19 Action Plan [NF/080 - INQ000280727]). Instead, the temporal profile of the epidemic used for the NHS planning assumptions was presented relative to the peak week of the epidemic. The only reference to time in the RWCS was that for an unmitigated epidemic, 95% of cases would be expected to occur in a 9-week period centred on the peak of the epidemic. Conversely, new RWCS scenarios agreed from April 2020 onwards did have an explicit representation of dates.

March 24th onwards

266. I will not detail every aspect of the scaling-up and broadening of surveillance following the announcement of the first lockdown; UKHSA and DHSC are in a better position to give a detailed timeline. I discuss genetic surveillance (COG-UK) in section K.

267. However, it is perhaps useful for me to summarise the value of the individual data-streams we made use of (other than genetic data, covered in section K), as they became available after 23rd March:

a. Case data – as of 23rd March, nearly all PCR testing capacity was committed to testing NHS hospital patients and individuals from clinically vulnerable groups – what later became Pillar 1 of the UK government's testing strategy. I, together with most academic participants on SAGE, pressed for testing to be scaled up to allow general community testing as soon as possible, together with the capability to implement large scale contact-tracing. I felt the wide-scale availability of testing would improve the effectiveness of case-isolation at preventing onwards transmission, and that contact-tracing might reduce the need for stringent NPIs to be maintained. Freely available PCR testing ("Pillar 2") started from 18th May 2020, later supplemented by lateral flow testing. We made use of Pillar 2 PCR test data (provided in the case linelist) in our real-time modelling for SPI-M-O from 1st June 2020.

b. CHESS and CO-CIN provided detailed data on individual hospitalised patients, including length of stay in general wards and ICU, the proportion of patients in different age groups

who were escalated to ICU, and clinical outcomes (death or discharge). These were critical parameters for the real-time modelling we undertook throughout the pandemic using the CovidSim (until April 2020) and then SirCovid models (April onwards). Neither dataset provided complete information on all hospitalised cases, but CO-CIN provided more representative data overall, since CHESS was somewhat biased towards ICU cases. It took until 26th March 2020 for us to gain access to CO-CIN.

c. SPI-M-O groups struggled to gain access to the definitive complete record of hospital episodes provided by the "SUS" dataset. The value of SUS is its completeness, and the ability to link individual hospital episodes to other datasets (e.g. COVID-19 testing or vaccination) via NHS number. This is needed to evaluate, for instance, the relative propensity of different COVID-19 variants to lead to disease severe enough to warrant hospitalisation. It was only in September 2021 that SPI-M-O finally got access to SUS-based records of hospital episodes associated with COVID.

d. Mortality data – data on lab-confirmed (i.e. PCR tested) COVID-19-associated deaths was provided in a dedicated deaths line list from 18th March 2020. This used the same anonymous identifier as the main linelist, allowing us to link the two datasets. It allowed us to track how hospital fatality rates varied over time, as well as providing another indicator of the overall progress of the epidemic. In 2021, it allowed us to characterise differences in lethality of different COVID-19 variants, and to estimate the effectiveness of vaccines at preventing COVID-19-related death. However, the death line list did not provide data on deaths suspected to be COVID-19 related but where no PCR test had been performed – which included a high proportion of deaths in care homes. SPI-M-O groups were therefore provided access to individual death data (based on death registrations) from May 2020 onwards (though we had access to the public summary data earlier). This listed all deaths where COVID-19 was listed as a contributory cause of death on the death certificate. It was not possible to link the ONS data to other individual-level datasets, but the ONS data provided a more comprehensive picture of the total mortality associated with COVID-19.

e. Serology data – *i.e.* data on antibody prevalence in the population – was invaluable in 2020 in providing data on the cumulative proportion of the population that had so far been infected in the pandemic. Among other things, this allowed us to derive more definitive estimates of the IFR. Prior to its availability in late April 2020, we had no

definitive way to estimate the number of people who had been infected, as compared with those (a small minority of infections prior to June 2020) who had tested PCR-positive for infection. The initial serology data we received in April was generated by PHE and was derived from so-called "convenience sampling" of blood donations. In July 2020, the REACT-2 core study run by Imperial College gave more definitive data based on random population sampling. Serology data needs careful analysis, correcting for the sensitivity and specificity of the tests used.

f. Infection prevalence – the tracking of infection prevalence (the proportion of the population with active COVID-19 infection detectable through a PCR test of a nasal swab) was an innovation in the COVID-19 pandemic, and a data source which was unique to the UK. As with serology, the motivation of setting up the ONS household survey and later the REACT-1 study was to provide unbiased data on the extent of infection in the population over time, recognising that only a proportion of infected individuals would ever seek testing (or prior to June 2020, be able to access a test). Again, these data were invaluable for the real-time modelling undertaken by ICCRT in allowing us to estimate true infection incidence as compared with just the incidence of symptomatic infection.

g. Vaccination data – from January 2021, SPI-M-O groups gained access to an anonymised version of the NIMS database which recorded every vaccination dose administered in England. These data could be linked to COVID-19 linelist data (cases, deaths, hospitalisations and variants), allowing vaccine effectiveness (VE) to be estimated, together with the impact of variants on VE. We used this data in the SirCovid model to reproduce past vaccination rates by age, region and day.

268. While access to every dataset wasn't always as timely as we would have preferred, the ability of SPI-M-O groups to link the different individual-level datasets together was invaluable, particularly as new variants of concern arose.

269. The two infection surveys also provided large-scale and more systematic data on the proportion of COVID-19 infections which cause symptomatic disease than many other studies (see para 124) – an evidence gap that Chris Whitty raised in an email exchange on 29th March 2020 [NF/234 - INQ000148826; NF/235 - INQ000148827; NF/236 - INQ000148828; NF/237 - INQ000148829; NF/238 - INQ000148830; NF/239 - INQ000148831]. In that email exchange he focussed on the potential value of serological surveys – which do give an invaluable

assessment of cumulative infection rates, but can only give assessments of the symptomatic fraction if the people who are serologically screened are also asked about their symptom history. More generally, estimates of the symptomatic fraction will depend on precisely what symptoms are included in the definition of symptomatic and such estimates are also complicated by the non-specific nature of many symptoms, which means many uninfected individuals show them. Hence having some non-specific symptoms (e.g. a mild fever) may or may not be associated with a contemporaneous infection. We accounted for these issues in an analysis of the symptomatic fraction based on REACT-1 survey data we undertook in early 2021 and published as part of our analysis of the first wave of the pandemic (Figure S7 [NF/011 - INQ000212077], see also para 124).

I. Timing and effectiveness of suppression

Note: with the agreement of the Inquiry, factual descriptions of activities undertaken by ICCRT provided in this section were written with the assistance of Prof Samir Bhatt and Drs Charles Whittaker, Daniel Laydon and Oliver Watson. All statements of opinion are entirely my own.

Introduction

270. Here I review ICCRT research to retrospectively evaluate the effectiveness of NPIs at reducing transmission, and comment on various aspects of NPI use during the pandemic. As previously, I use the term suppression to refer to a set of interventions (typically NPIs) or behavioural recommendations implemented to control the spread sufficiently to bring R to below 1, the threshold for epidemic growth. Such measures are intended to reduce the contact rate between infected and susceptible individuals and/or the probability of an infectious individual successfully transmitting the virus.

271. Population contact rates can reduce spontaneously through risk avoidance behaviour in the population, or as the result of mandatory NPIs or public health recommendations made by governments or public health agencies. As discussed in section G, NPIs span a wide range of measures, with community-focussed measures including physical/social-distancing, bans

on gatherings above a certain size involving people from multiple households, closure of hospitality and entertainment venues, working from home and school and university closure.

Interventions leading to epidemic suppression

272. In March 2020, nearly all European countries introduced interventions which achieved suppression, even if in some instances that was not the stated goal. There was considerable variation between countries in the exact combination of measures implemented, their relative timing, and the overall degree of stringency.

273. Even when the same term was used to describe a policy, there were large differences in the implementation details between different countries. In particular, the term "lockdown" did not have a consistent definition during the pandemic. Instead, it was used to describe varying sets of policies (including banning of gatherings above a given size, closure of businesses, and stay or work at home orders) aimed at suppressing transmission. As an example, whilst the measures constituting the Danish lockdown banned public gatherings of over ten individuals, the first UK lockdown measures banned public gatherings of over two individuals. This variation was complicated further by subsequent refinement of policies over time. For example, the first and third lockdowns in England involved closure of schools, while the second did not.

274. Sweden is frequently cited as a country which did not "lockdown" but that successfully achieved suppression. Rather than imposing mandatory measures, guidelines were issued by the government recommending social distancing, avoiding non-essential travel, and working from home when possible. Gatherings of more than 50 people were banned, secondary schools and universities switched to remote learning but most businesses and primary schools largely remained open. In a study we undertook comparing the first wave of the pandemic in the UK, Denmark and Sweden [NF/022 - INQ000262573], we found that Sweden achieved almost the same level of reduction in contact rates (as reflected in workplace mobility changes and the proportion of people avoiding public spaces) as many other European countries, including its neighbour Denmark which instituted a suite of mandatory policies [NF/022 - INQ000262573]. However Swedish control measures took longer to reduce R below 1 than did measures adopted by both the UK and Denmark, and achieved a lower overall reduction in R.

275. I therefore reject the criticism that in the ICCRT Report 12 [NF/002 - INQ000262593] we "predicted" that 42,000 deaths would occur in Sweden if that country didn't lockdown, which is based on falsely equating suppression (what we modelled) with lockdown, and mitigation with the policies adopted in Sweden. I also note that Report 12 does not use the word "lockdown" once.

276. I also note that, in the absence of sufficiently intense NPIs, other countries failed to achieve suppression. An example of this is Brazil, which did not institute a national lockdown (though some constituent states did adopt comparable measures), but where, unlike Sweden, suppression was not achieved until the autumn of 2020, and then only likely in part due to the substantial accumulation of immunity in the population by that time [NF/240 - INQ000262613].

277. We will never know whether a more voluntary approach might have achieved suppression in the UK context, and given the cultural determinants of behaviour, it is not a question that epidemiology can readily address. However, whether due to mandatory or voluntary NPIs, achieving suppression would still have required a reduction in contact rates similar to that which was achieved in reality: such reductions would inevitably involve a major departure from normal day-to-day behaviours.

278. From our later work analysing the impact on transmission of the "Tier" system of local COVID-19 alert levels and NPIs adopted in England in October 2020 [NF/241 - INQ000262598] (see also section J) we can speculate that suppression might have been achieved with largely voluntary recommendations in some local areas of the UK but not in others. However, by October 2020, considerable effort had been invested in making the UK "COVID-safe", via a combination of organisationally-focussed safety measures and individual protective behaviours intended to reduce the chances of transmission in public venues. Such measures were not in place in March 2020.

Timing of suppression measures

279. In the context of a rapidly growing epidemic, both the timing and effectiveness of suppression measures can have disproportionately large effects on total mortality seen in that wave of transmission (*i.e.* until the policies are lifted) [NF/011 - INQ000212077; NF/022 - INQ000262573; NF/242 - INQ000262580; NF/243 - INQ000262589]. Earlier responses reduce the total number of people infected in the epidemic until suppression drives case numbers to

low levels, and therefore reduces mortality and hospital admissions. In an epidemic which is growing exponentially, there is correspondingly an exponential relationship between the number of infections and deaths and the time delay before interventions are introduced.

280. In evidence to a House of Commons Select committee on 10th June 2020, I stated that had interventions been introduced a week earlier in the UK, the number of deaths up to that date might have been halved. This was a deliberately conservative statement, given our best estimate of the doubling time of the epidemic in mid-March 2020 was around 4 days. In later work, we explicitly modelled the counterfactual scenario of moving the lockdown of 23rd March back to 16th March [NF/011 - INQ000212077], and estimated mortality up to the end of November 2020 would have been reduced by 48%; the reason the estimated reduction wasn't substantially larger is that the model accounted for the autumn 2020 wave growing slightly faster had infection levels been kept lower in the spring.

281. Relatedly, interventions which cause a larger decrease in R (*i.e.* are more effective at reducing transmission) lead to a more rapid decline in infection incidence, fewer individuals infected in total, and therefore fewer deaths. For instance, while the measures adopted by Sweden and Denmark both achieved suppression, the proportional reduction in R was larger in Denmark [NF/022 - INQ000262573].

282. An important corollary of these points is that there is a trade-off between the timing and effectiveness of suppression measures. Acting earlier allows one to introduce somewhat less stringent measures (so long as R is brought below 1) at no greater cost in terms of total deaths than acting later with more stringent (and effective) measures.

283. Differences in the timing (relative to the stage of the epidemic) and effectiveness of NPIs introduced in March 2020 explain much but not all of the variation in first wave mortality seen across European countries [NF/243 - INQ000262589]. Other factors include population demography, prevalence of comorbidities, healthcare system capacity and the extent of infection in care homes.

284. How COVID-19 mortality was measured also varied – countries differed substantially in the extent of testing, especially in the first wave, and in the criteria adopted to define a COVID-associated death. For retrospective analyses, excess deaths estimates provide the most objective measure of the mortality impact of the pandemic, especially for countries with complete death registration.

285. Spontaneous/voluntary behavioural changes may have also been a potentially important contributor to reductions in R throughout the pandemic. Results from the CoMix social contacts survey [NF/244 - INQ000223808; NF/245 - INQ000092690] demonstrate a substantial lag between relaxation of interventions and the return of contact rates to levels seen prior to imposition of controls.

Estimating the effectiveness of NPIs

286. Over the course of the pandemic, the team at Imperial College undertook a substantial amount of work to estimate how effective different NPIs had been at reducing transmission (*i.e. R*). There are two key challenges that arise when attempting to estimate NPI effectiveness. The first is accounting for the delays between changes in transmission, and the resulting changes in COVID-19 cases and/or deaths. Specifically, an observed reduction in deaths today is the result of changes in transmission sometime in the past, which is not directly observed. To address this challenge, and capture the dynamics required to infer NPI effectiveness, the Imperial College team developed a statistical inference model (Epidemia – see section E) based on a mathematical process called the renewal equation [NF/019 - INQ000228162] that was used to infer both temporal changes in R and the comparative impact of different NPIs on R from COVID-19 death data and details of the timing of different NPIs.

287. A second challenge arises from the fact that in each country, multiple control measures were often imposed on the same day (especially in March 2020). When considering a single country alone, this prevents attribution of changes in transmission to a particular control measure. This issue can be mitigated however by leveraging the fact that different countries implemented different combinations of control measures, at different times and on different days – this variation between countries in which and when control measures were implemented enables the effect of an individual control measure to be inferred.

288. We published an initial analysis of the impacts of governmental interventions in 11 European countries in Report 13 [NF/018 - INQ000262556], subsequently published in Nature [NF/020 - INQ000262594]. Subsequent work substantially refined this analysis. Improvements included an expansion of the number of European companies included [NF/020 - INQ000262594] and a more granular representation of different classes of control measures, including decomposition of "lockdown" into its constituent measures [NF/020 -

INQ000262594; NF/024 - INQ000218211], and a comparison of the effectiveness of NPIs in the first and second waves of the pandemic in Europe [NF/024 - INQ000218211].

J. March 23rd 2020 to January 2021: epidemic trends, modelling and policy responses

Note: with the agreement of the Inquiry, factual descriptions of activities undertaken by ICCRT provided in this section were written with the assistance of Drs Anne Cori, Lilith Whittles, Natsuko Imai and Marc Baguelin. All statements of opinion are entirely my own.

24th March-31st August 2020

289. Our input into SPI-M-O and SAGE meetings between 24th and 31st March 2020 focussed on developing a range of planning scenarios for the trajectory of the UK epidemic for use within the NHS and UK government. This culminated in a new RWC planning scenario [NF/246 - INQ000233787] being adopted on 29th March [NF/247 -INQ000221758]. It should be emphasised that this work was not intended to predict the trajectory of the epidemic. In retrospect, this RWC fell below peak weekly deaths and hospitalisation in the first wave but exceeded (as intended) cumulative deaths and hospitalisations to 1st September. There were a number of reasons for the RWC undershooting the peak: (a) modelling was undertaken before the peak had been reached, and in retrospect over-estimated the impact of measures taken on 16th March, hence predicting infections would peak earlier than observed; (b) estimates of hospital stay duration (in general beds and ICUs) and hospital case fatality ratios were preliminary at that stage, and were only refined in April as much more detailed hospitalisation data became available (see section H).

290. From April 2020 onwards, SPI-M-O developed a more formal process for generating short- and medium-term projections of the epidemic trajectory and estimates of R. This was based on averaging across the results from multiple models developed by the groups participating in SPI-M-O (a so-called model ensemble approach). As we finalised development of the (less computationally intensive) SirCovid stochastic compartmental model (see Section E), we initially generated projections and R estimates using the CovidSim simulation model used for Report 9 and previous modelling of NPIs; however, that became increasingly computationally untenable over time – large scale individual-based simulations are not optimally suited to being fitted to epidemic data in real-time. We therefore switched from using CovidSim to

SirCovid for most work on the UK epidemic from 16th April 2020 , which was the date we first used SirCovid to generate estimates of R and short term projections. SirCovid had the advantage of being designed from scratch to be fitted in real-time (and routinely) to the surveillance data streams being collected in the UK.

291. The one other major topic where we provided modelling of NPI intervention options to SAGE prior to September 2020 was the modelling of exit strategies from the first lockdown, commissioned by SAGE. This work still made use of the CovidSim model, which was better suited to modelling complex NPIs, particularly contact tracing. Multiple SPI-M-O groups undertook this work. Preliminary analyses were presented at the 21st April [NF/248 - INQ000074997] and 28th April [NF/249 - INQ000262607] SAGE meetings, and finalised analyses (following a schedule for phased lifting of measures provided to SPI-M-O groups by the Cabinet Office) to the 5th May SAGE meeting [NF/033 - INQ000280728]. The effect of Test and Trace was included in this modelling, making (in retrospect) optimistic assumptions about the effectiveness of contact tracing. Our paper concluded "All scenarios are predicted to lead to large autumn waves of transmission, healthcare demand and deaths in the absence of additional control measures or higher long-term levels of social distancing". We additionally included the potential impact of reactive transient closure of schools and workplaces in response to cases being detected, but concluded that in the autumn "a 50% reduction in potentially infectious social contacts outside schools and workplaces is needed to maintain R<1, even with contact tracing and reactive school and workplace closure". The 5th May SAGE meeting [NF/250 - INQ000120512] concluded that relaxation of NPIs should be based on assessment of incidence levels and other relevant data (e.g. R) rather than predetermined dates.

292. This exit strategy modelling reconsidered age-stratified social distancing (*i.e.* enhanced shielding) and concluded that even if 80% effective (likely above what was feasible, given the level of social care and health care support required by the most COVID-vulnerable), such a strategy would still lead to overwhelming healthcare demand in the absence of NPIs applied to the rest of the population.

293. I also contributed to a more conceptual paper ("Post lockdown epidemiological scenarios"), co-authored by Angela McLean, John Edmunds, Graham Medley and myself (with later input from Patrick Vallance) which evaluated the pluses and minuses of running the epidemic "hot" (high infection levels, but within NHS capacity limits) or "cold" (keeping

infections at low levels). This paper, together with ICCRT exit strategy modelling discussed above, and both LSHTM and ICCRT modelling of testing requirements for an effective test and trace system, were discussed at two "post lockdown discussion" online meetings organised by GoS on 24th and 27th April, attended by Chris Whitty, Patrick Vallance, Angela McLean, Graham Medley, John Edmunds and myself (plus some other officials). These discussions built on an earlier small group meeting I include the relevant meeting materials [NF/251 - INQ000148832; NF/252 - INQ000148833; NF/253 - INQ000148834; NF/254 - INQ000148835; NF/255 - INQ000148836; NF/256 - INQ000148837; NF/257 - INQ000148838; NF/258 - INQ000148839; NF/259 - INQ000148840; NF/260 - INQ000148841; NF/261 - INQ000148842; NF/262 - INQ000148845; NF/263 - INQ000148846; NF/264 - INQ000148847; NF/265 - INQ000148848; NF/266 INQ000148849; NF/267 - INQ000148850]. From memory, there was general consensus that the preferred objective would be to keep infection levels low. I am not sure what the follow-up to that discussion was within government, though from the May 10th COVID-19 Alert level strategy [NF/268 - INQ000065338] it appears a high infection level strategy was adopted.

294. In parallel, Nick Grassly, a professor in MRC-GIDA led on a more detailed analysis of the role of testing (both PCR and antibody) in controlling transmission, published initially as our Report 16 [NF/005 - INQ000262557000000], and later as a journal article [NF/004 - INQ000215948]. While not submitted as a SAGE document, this work was influential in informing scaling up of testing and later use of mass testing, and Prof Grassly attended a number of ad hoc meetings with civil servants (I think from DHSC and Cabinet Office) after its publication.

295. We also submitted an analysis of the intervention policies adopted by South Korea to the 5th May SAGE meeting [NF/269 - INQ000229143] which emphasised that case cluster investigations were perhaps more important in that country's success in controlling spread than individual contact tracing.

Stepping back from SAGE

296. Immediately after the May 5th SAGE meeting, I was contacted by the Daily Telegraph regarding a breach of lockdown rules. Patrick Vallance and I agreed that I should "step back" from SAGE and I issued a statement to this effect, deeply regretting the breach and its potential impact. I take this opportunity to repeat the apology which I have made previously.

I therefore no longer attended SAGE meetings from May 6th 2020 until January 2022, though I retained access to SAGE documents. As an ad-hoc group rather than a statutory committee, SAGE has participants (invited by the co-chairs) rather than formal membership, so I continued to be listed as a participant in SAGE-COVID throughout the pandemic. After discussion with Patrick Vallance and Angela McLean, I continued to attend meetings of other advisory committees and groups, on the basis that SAGE itself was the only group that directly advised ministers on policy, and that the more science-focused groups feeding into SAGE would benefit from me continuing to represent one of the largest UK epidemiological research centres. This is referred to in the minutes of the 7th May SAGE meeting [NF/270 - INQ000120513]. No other committee chairs or officials raised this issue with me. The only other SAGE meeting I attended was on 7th January 2022, to discuss my ongoing analysis of the severity of the Omicron variant.

297. I attended a number of small group meetings with Patrick Vallance and/or Chris Whitty between May 2020 and March 2021. Reviewing my calendar, I can find entries for 29th May 2020 [NF/271 - INQ000062300] – I believe concerning Test and Trace and data flows), 9th June 2020, 24th July, 11th August 2020 (with Dido Harding, concerning the effectiveness of Test and Trace), 7th September 2020, 7th October 2020, 14th October 2020, 20th November 2020, 14th December 2020 (about the Alpha variant), 18th December 2020 (about Alpha), 21st December 2020 (about Alpha), 8th January 2021 (Alpha and horizon-scanning) and 18th March 2021. Most (other than the ones concerning the Alpha variant) were scientifically-focussed horizon-scanning meetings and as such did not have agendas or minutes. I additionally participated in more formal meetings reviewing the performance of SAGE (25th June 2020 and 15th October 2021). I also had a number of one-on-one informal conversations with Patrick Vallance over the course of the pandemic. I did not keep records of any of the small group meetings (or other informal conversations), and in the absence of minutes, have little recollection of the details of what were often wide-ranging discussions. An exception to this is those concerning Alpha, as those meetings were solely focussed on that variant's epidemiological characteristics and the consequence for epidemic growth (section K).

298. From May 2020 onwards, while ICCRT continued to actively support SPI-M-O via the provision of modelling in response to commissions from SAGE or the SPI-M-O secretariat (e.g. weekly projections and R estimates and modelling of intervention options), we only undertook scenario modelling of prospective UK policy options in response to such

commissions. Reflecting evidence gaps and the requirements of global partners, we also increased the priority given to analyses of epidemiological data to estimate impacts and parameters (e.g. excess mortality, hospital demand, viral evolution, IFR) and to non-UK focussed modelling.

COVID-19 Alert system, May 2020

299. The Prime Minister announced a strategy focussed around a hierarchy of five COVID-19 Alert levels on 10 May 2020 [NF/272 -INQ000273791; NF/273 -INQ000086784]. SAGE was not consulted on the strategy, to my knowledge, though it was consulted later on the criteria proposed by JBC to be used for moving between alert levels [NF/274 - INQ000119952; NF/275 - INQ000119951; NF/276 - INQ000262626]. It should be noted that the main emphasis at that time was on how NPIs could be safely de-escalated, rather than on the more politically challenging issue of escalation. My own view of the criteria proposed was that the de-escalation criteria seemed reasonable, but that emphasis on case incidence rather than growth rate or R (while justified in terms of the ability of Test and Trace to cope) risked escalation being delayed until infection levels were already very high. In addition, the strategy initially only envisaged a tightening of NPIs if there was a "material risk of healthcare services being overwhelmed" [NF/272 - INQ000273791].

300. I also felt this strategy used an inappropriately short time horizon to evaluate the costs of intervening. In doing so, it failed to learn perhaps the most important lesson from March 2020: namely that acting early saves lives and costs no more economically than acting late – because if government acts earlier, it can also lift measures earlier.

Updated RWCS, July 2020

301. In July 2020, the UK Cabinet Office requested SPI-M-O to provide a Reasonable Worst Case Scenario (RWCS) for government planning for winter 2020-21. This was to assume high, but plausible levels of transmission. The agreed scenarios of the commission stipulated the trajectory of incidence until the end of November, then followed by three different NPI scenarios (reduction of all non-household contacts to 30%, 40% and 50% of their normal level).

302. On 17th July, preliminary results from Imperial based on fitting our SirCovid model to the epidemic to date were presented to the SPI-M-O secretariat and Cabinet Office representatives alongside results from three other groups (University of Warwick, Faculty Ltd

and University of Exeter). On 21st July, a follow up meeting was held where it was collectively decided that the SirCovid model was best suited to respond to the RWCS commission. A report describing methodology and initial results for 6 scenarios was also shared with SPI-M that day.

303. Between 22nd July – 24th August, we sent the SPI-M-O secretariat output files covering a total of 18 scenarios. Each file gave projected hospital admissions, occupancy, and deaths, broken down by region and age. All scenarios considered assumed transmission levels would remain substantially lower than before the first lockdown was implemented, due to the implementation of "COVID-safe" protocols and behavioural caution in the population. As well as the RWCS scenario, we modelled a "laissez-faire scenario" which assumed R would remain at the high level assumed for late November (*i.e.* no measures implemented).

304. Our scenarios suggested that if transmission increased to the high but plausible values assumed in the RWCS, measures would need to be put in place in December 2020 to avoid a substantial second wave of hospitalisation and deaths similar in scale to the first wave. Failure to reduce R to below 1 from December onwards was projected to potentially result in over 64,000 deaths in hospitals between 1st September 2020 and 31st March 2021 and daily deaths remaining high (of the order of 500 per day) at the end of March 2021. I note that the UK saw 85,000 COVID-19 associated deaths in that period, though part of the reason this number is larger than the RWCS was because of the higher transmissibility and severity of the Alpha variant detected in December 2020.

305. Even under a more optimistic scenarios assuming a December reduction in non-household contacts to 35% of their normal pre-lockdown level, hospital bed occupancy was projected to remain high over the winter wave with ICU peak levels about half the peak of the first wave but sustained over a much longer period.

A growing epidemic, September 2020

306. I preface my detailed comments by noting that over two thirds of the UK's COVID-associated deaths reported prior to 1st April 2021 occurred after 1st September 2020. It is my view that a large majority of these latter deaths were avoidable, at no greater economic cost to the UK, had decisions to escalate NPIs been made more promptly in the autumn of 2020.

307. On 25th July 2020, I emailed Angela McLean, Graham Medley, John Edmunds and Patrick Vallance with a simple comparison of REACT-2 seroprevalence by age, pillar 2 case incidence by age and typical (pre-pandemic) contact rates by age – for adults 18 and over [NF/277 -

INQ000148975; NF/278 - INQ000148976; NF/279 - INQ000148977; NF/280 - INQ000148978; NF/281 - INQ000148979; NF/282 - INQ000148980; NF/283 -INQ000148981]. My intention was to highlight that the pre-pandemic contact rate surveys (used by all SPI-M-O models) had predicted infection attack rates in the first wave of the pandemic well. I also noted that there was little evidence, comparing symptomatic pillar 2 case incidence with seroprevalence (which measures underlying infections), that the proportion of infections which were symptomatic varied much by age (among adults), except perhaps in those over 75 (where case rates were lower than might be expected). This analysis was among the papers for the 29th September SPI-M-O meeting. More significantly, the same email discussion also noted that there was a clear signal of increasing Pillar 2 cases by that time, across all ages. John Edmunds also commented that the LSHTM CoMix survey was recording increases in population contact rates at that time, especially among working age adults.

308. The first SPI-M-O Consensus statement to definitively (i.*e.* with greater than 95% statistical certainty) conclude that R for the UK as a whole was above one was that of 10th September [NF/284 - INQ000262623], tabled at the SAGE meeting on the same date. The gap from 25th July to 10th September reflects that SPI-M-O R estimates were based on multiple indicators – cases, hospitalisation and deaths – all of which had different lags in representing underlying trends in transmission, In addition, at least two weeks of increases in any one indicator were needed to definitively resolve an increase. Given these lags, this suggested the UK epidemic started growing again in August 2020, and likely earlier in some regions.

309. The increasing levels of transmission seen at that time were entirely to be expected, given the relaxation in NPIs over the previous 4 months. Similar trends were seen in most if not all other European countries. However, growth rates were substantially lower than seen in March 2020, reflecting continuing NPIs, the impact of Test and Trace, the investment in making public spaces "COVID-safe", and continued risk averse behaviour by the population.

310. On 20th September, the Prime Minister met Sunetra Gupta, Carl Heneghan and Anders Tegnell, to apparently hear alternative scientific perspectives from that provided by SAGE. Angela McLean and John Edmunds represented SAGE, and Patrick Vallance and Chris Whitty were present. I will leave the attendees to comment on what was discussed, though I was pleased to learn afterwards that SAGE's authority had if anything been reinforced by the meeting.

311. A day later, on 21st September 2020, the UK government followed the strategy set out in its 10th May COVID-19 alert levels document [NF/268 - INQ000065338], and raised the COVID-19 alert level to 4. On the same date, SAGE concluded that a substantive package of new NPIs would be needed to slow or reverse the exponential growth in the epidemic then being seen, and once more explained the benefits of acting early [NF/285 - INQ000061566]. I agreed with these conclusions. I had been in discussion with Patrick Vallance, John Edmunds and Graham Medley concerning what measures would be needed to slow or reverse epidemic growth over the week prior to that meeting. I cowrote a first draft of the table of NPIs presented at that SAGE meeting, with John Edmunds and Matthew (Matt) Keeling [NF/286 - INQ000231400].

312. On 22nd September, the Prime Minister announced a package of NPIs [NF/287 - INQ000237538], the most significant of which were a recommendation to work from home where possible, a tightening of restrictions around hospitality venues and an expansion of mask wearing. I felt that while the measures announced were certainly better than taking no action, they did not reach the level recommended by SAGE the previous day and would at best slow the rate of growth of the epidemic.

313. The SPI-M-O consensus statement considered at the 1st October SAGE meeting noted that 5 SPI-M-O models estimated that the RWCS would be breached in the following two weeks [NF/288 - INQ000262622]. This conclusion was reflected in the minutes of that SAGE meeting [NF/289 -INQ000120560].

314. On 9th October 2020, SPI-M-O requested further work on potential scenarios for winter 2020-21. We, alongside two other academic groups (LSHTM & Warwick) and PHE, produced scenarios assuming different transmission rates and effectiveness of NPIs. We shared the results of this scenario modelling with SPI-M-O and the results were used by the SPI-M-O secretariat to produce slides that were tabled at the SAGE meeting on 15th October.

315. Also on 15th October, we shared with SPI-M-O a report titled "Potential impact of a circuit breaker" [NF/290 - INQ000148968]. This report concluded that implementing a two-week "circuit-breaker" lockdown with the same level of effectiveness as the first lockdown might slow down the UK epidemic by up to 26 days compared with the no "circuit-breaker" scenario. It also showed that the effectiveness of the intervention would heavily rely on compliance, as achieving only only half the reduction in transmission achieved during the first lockdown would result in only a 10-day delay.

The "Tiers" system

316. From July 2020, piecemeal additional NPIs had been introduced in "hotspot" areas of England with particularly high and growing case incidence: Leicester on 4th July, Luton and Blackburn with Darwen on 25th July, Bradford on 1st August, followed by wider restrictions across several areas of the North of England on 5th August. The latter were adjusted multiple times in the following six weeks, both in geographic scope and the precise restrictions applied. On 14th October, these piecemeal restrictions were replaced (in England) with the more systematic system of local COVID-19 Alert levels (otherwise called Tiers). Local area alert levels were set as medium (Tier 1), high (Tier 2) or very high (Tier 3) based on a never very transparent evaluation of local growth rates (or R), pillar 2 case incidence per 100,000 of population and hospital admissions/occupancy. NPI stringency increased with the level of alert; with Tier 3 banning all indoor mixing between different households, and closing several types of leisure venues (*e.g.* gyms).

317. To my knowledge, SAGE and SPI-M-O were not consulted ahead of the introduction of the local Tier system, and were not asked to review the specific criteria used by DHSC to determine local alert levels.

318. My own view is that while the Tier system was a logical extension of the May national COVID-19 alert level system, it suffered from the same conceptual flaw in implementation; namely, that escalation to the "very high" level (Tier 3) seemed to be restricted to areas of highest case incidence and healthcare demand. I felt that since only Tier 3 measures were likely to make a substantial difference to transmission, the result of the Tier system would be a gradual "levelling-up" of infection rates to high levels across the country.

319. We used the ICCRT Epidemia model (at that time being used to publish weekly estimates of R and short term projections for every LTLA in the UK) to estimate the impact of Tiers on transmission [NF/244 - INQ000262598], sharing preliminary results with SPI-M-O on November 4th 2020, and an updated version thereafter. In our finalised analysis, we estimated Tiers 1 and 2 reduced R by less than 7%, but that Tier 3 achieved a 23% reduction (compared with what R would have been without Tier 3 restrictions). Had all local authorities been moved to Tier 3 (effectively a "lockdown-lite") by the end of October, we estimated the Tier 3 measures would have suppressed transmission (*i.e.* achieved R<1) in 93% of local authority areas.

The second England lockdown

320. On October 31st, the Prime Minister announced a one month second lockdown would start in England on November 5th 2020. My understanding is that this rather sudden policy pivot resulted from an intense set of interactions between Number 10 officials and key SAGE participants on October 30th. I was not involved in these and do not know further details or who precisely was involved, as I was taking a rare day off on the 30th. Hence I declined an invitation from Jeremy Farrar to a call early that afternoon.

321. For the reasons I have explained above, I fully agreed with the decision to introduce the second lockdown. However, I felt uncomfortable with aspects of how SPI-M-O modelling results were used as part of the evidence supporting and/or justifying that decision.

322. A slide containing modelling results was leaked to the BBC on 31st October [NF/291 - INQ000262576] which purported to show "forecasts" of the epidemic trajectory for the next few months from 4 modelling groups (including ICCRT). In fact, these were modelling outputs produced over two weeks earlier by the four groups in response to the SPI-M-O request for "reasonably bad" winter scenario modelling (para 314 above). The modelling pre-dated the Tier system and therefore did not account for its effects. This leaked modelling was then explained post hoc by a SPI-M-O paper published as an addendum to the 28th October 2020 SAGE meeting archive [NF/292 - INQ000262555]. However, that modelling was never considered by SAGE before or after the decision to move to a second lockdown.

323. Given these issues, I felt the apparent use of that modelling to influence the decision to move to a second lockdown was problematic. The leak also further added to criticism of SPI-M-O modelling in some sections of the media.

324. It would have been much preferable for decision-making to have relied solely on the latest SPI-M-O medium term projections [NF/293 - INQ000262625] which had been considered by the 28th October SAGE meeting and which still showed hospital admissions exceeding April 2020 levels by mid-November in the absence of the introduction of more stringent national NPIs. Indeed, it was this modelling that Patrick Vallance largely focussed on in the press conference of 31st October 2020 [NF/294 - INQ000262579]. However, I do not know whether those medium term projections were presented and discussed with ministers on 30th October.

325. The second national lockdown in England was less stringent than the first. Notably, schools remained opened, and the definition of essential workers was broader.

326. On November 3rd, we shared with SPI-M-O a report on the "Potential impact of a second national lockdown on the COVID-19 epidemic in England" [NF/295 - INQ000148865] [210] This was presented and discussed at the SPI-M-O meeting on 4th November. The results suggested that it would be unlikely that the second lockdown would slow epidemic growth sufficiently to prevent a peak in hospital admissions and deaths higher than that seen in the first wave later in the winter. We also looked at various scenarios around loosening restrictions over Christmas, all of which further increased transmission. To avoid a major winter wave larger than the first wave of the pandemic, our modelling suggested that additional measures would be necessary following the second lockdown, including a more stringent tier system and potentially a third lockdown in January 2021.

327. From November 2020, SPI-M-O regularly included additional simple counterfactual scenarios on top of the regular medium-term projections. These additional scenarios assumed levels of transmission that were slightly lower and slightly higher than current latest estimates. This was intended to provide a reference point for policy makers to help interpret the potential impacts of introducing or lifting interventions (e.g. lifting NPIs and introducing vaccination).

328. The second lockdown caused the epidemic to slightly decline in England overall. However, case numbers and hospitalisations accelerated sharply in some regions (East of England, London and the South East) after the lockdown ended on 2nd December.

Alpha and the third England lockdown

329. A revised, more stringent version of the Tier system was introduced at the end of the second lockdown, with many more local authority areas being classified as Tier 2 or 3 than had been the case on 5th November. As case incidence and hospitalisations grew in the first two weeks of December, many more areas in London, the East of England and the South East were escalated to Tier 3 by 14th December.

330. By the week of 14th December, the potential role of the B.1.1.7 (Alpha) variant of concern in driving the growth rate of the epidemic was starting to be understood. The role of ICCRT in identifying and characterising the Alpha variant is discussed in more detail in section K. However, the discovery of Alpha rapidly led to the introduction of a fourth tier of the local alert level system – with restrictions very similar to the second lockdown – on 19th December [NF/296 - INQ000086621].

331. On 4th January 2021, the prime minister announced the third England lockdown [NF/297 - INQ000065415]. Due to concerns about the increased transmissibility of the Alpha variant, the measures introduced were more stringent than for the second lockdown, and included closure of all schools. The latter decision was based on an assessment not just of an increase in school-based outbreaks associated with the Alpha variant, but also because of concerns about incidental transmission which might occur due to schools being open (associated with school drop-off and pick-up, for instance).

332. We did not undertake prospective modelling of the impact of a third lockdown on the Alpha variant ahead of it being announced. In large part this was a prioritisation decision; other SPI-M-O groups were doing so, and between 11th and 31st December 2020 we were focussed more on characterising the transmission advantage of Alpha relative to pre-Alpha strains of virus in circulation [NF/298 - INQ000230152]. This work is discussed more fully in section K.

333. In November and December 2020, we undertook modelling of the potential impact of vaccination roll-out on the pandemic in 2021 which I discuss in section L.

Effectiveness of Test and Trace

334. I thought it would be useful for me to summarise my views on and involvement in evaluating the effectiveness of the Test and Trace system in England. While the topic spans the period June 2020 to the end of the pandemic, I include it here because Test and Trace was launched in late May 2020.

335. Test and Trace had three components:

a. Free, rapid testing of suspect COVID-19 cases in the community (Pillar 2 testing). Initially only PCR-based testing (via self-swabbed samples transported by post), but lateral flow tests for institutional and later home use were rolled out from December 2020. PCR testing volumes expanded from approximately 60 thousand tests per day in June 2020 to 600 thousand per day in December 2021.

b. Contact tracing. Individuals testing positive for COVID-19 via pillar 2 were contacted by phone, text and/or email requesting details of recent close contacts. This data was either recorded as part of a phone interview, or via the Test and Trace website.

c. The NHS COVID-19 smartphone app. This was launched in September 2020 and used low energy Bluetooth signals to track close-proximity interactions with other app users. When

an app user tested positive for COVID-19, the app infrastructure could then notify contacts that they had potentially been exposed to infection, requesting that they self-isolate. Due to technical issues and privacy concerns, the app was entirely anonymous, so users categorised as contacts of a case by the app were not able to be personally identified by Test and Trace unless they informed Test and Trace of having received an app notification.

336. I had been involved in SAGE discussions and work in April 2020 examining the potential impact of a Test and Trace system and the capacity for contact tracing an effective system would require. These issues were discussed at the April 24th and 27th meetings mentioned in paragraph 293 above [NF/251 - INQ000148832; NF/252 - INQ000148833; NF/253 - INQ000148834; NF/254 - INQ000148835; NF/255 - INQ000148836; NF/256 - INQ000148837; NF/257 - INQ000148838; NF/258 - INQ000148839; NF/259 - INQ000148840; NF/260 - INQ000148841; NF/261 - INQ000148842; NF/262 - INQ000148845; NF/263 - INQ000148846; NF/264 - INQ000148847; NF/265 - INQ000148848; NF/266 - INQ000148849; NF/267 - INQ000148850]. Both our modelling and that of LSHTM suggested that (outside of a lockdown context) 10-30 contacts per index case would need to be contacted and to self-isolate for Test and Trace to have a major impact on transmission (*i.e.* to reduce R by 40% or more). In addition, contact tracing needed to be rapid to have a substantial impact – all contacts being notified within 48-72h of symptom onset in the index case.

337. Test and Trace never came close to achieving these contact tracing goals; excluding contacts identified via the NHS COVID app and via school based contact tracing, the average number of contacts reached per positive case rarely exceeded 3 (and was typically lower), and 90% of the contacts reached were household contacts, who, under household isolation rules, had to isolate in any case after a case in the household was identified.

338. I advised Test and Trace on methods to evaluate its impact on transmission at various points during the pandemic, including during with an informal meeting with Dido Harding and Patrick Vallance on 11th August 2020, and then later in 2021 in a sequence of advisory meetings which led to the publication of UKHSA's analysis of the effectiveness of Test and Trace in September 2021 [NF/299 - INQ000262568].

339. The UKHSA analysis concluded that Test and Trace may have reduced transmission by 20% between August 2020 and April 2021 (the study cut-off date). This corresponded to a reduction of approximately 0.2 in R. However, the study considered a counterfactual scenario

of Test and Trace not existing and estimated what the impact would have been of just continuing the March 2020 policy of symptomatic case isolation and associated household quarantine. Under that counterfactual, those policies were estimated to give an approximately 15% reduction in transmission, meaning Test and Trace only added another 5%. This small impact largely reflects: the small proportion of non-household contacts identified or reached; the delays in testing cases, identifying contacts and then reaching those contacts; that only approximately a third of COVID-19 infections in the UK between June 2020 and January 2022 were ever confirmed with PCR testing.

340. The UKHSA analysis did not attempt to quantify what I felt were the significant social benefits of the Test component of Test and Trace: the assurance provided to people with symptoms to have infection confirmed (or not); the likely higher compliance with self-isolation associated with a positive test results; the facilitation of risk avoidance (e.g. in preventing exposure of elderly relatives) provided by the ready availability of home (lateral flow) testing. Hence while I feel the non-app Trace component of Test and Trace failed to provide value for money, I think assessing whether Pillar 2 testing (which made up the bulk of the cost of Test and Trace) was worth the many billions of pounds invested is a much more complex undertaking.

341. I will not comment on evaluations of the NHS COVID app as I was not significantly involved in these, but retrospective analyses suggest it was somewhat more effective than the non-app-based contact tracing system run by Test and Trace, perhaps particularly at certain points of the pandemic (e.g. during the "pingdemic" of June-July 2021). I would recommend the Inquiry contact Christophe Fraser at the University of Oxford for an authoritative view on the development and effectiveness of the NHS COVID app.

Inequity in the impacts of COVID-19 and NPIs

342. The potential for the impacts of both the pandemic and NPIs to be inequitably distributed across society was appreciated by SAGE from February 2020 onwards, and I believe was frequently discussed on SPI-B in particular. The possible effect of income on the ability to comply with NPIs was discussed at the 3rd March SAGE meeting, with a recommendation that government address this potential issue [NF/075 - INQ000119719]. Care homes were specifically highlighted as a priority on March 10th, as were the likely challenges for households on low incomes [NF/124 - INQ000109125]. The March 13th meeting again discussed the challenges of preventing inequitable impacts of NPIs [NF/079 - INQ000109142].

Mitigating inequitable impacts of NPIs on low income population groups was also discussed on 26th March [NF/300 - INQ000119726].

343. Differences in UK COVID-19 mortality between ethnic groups began to be recognised from early April, once sufficient data accumulated. PHE was commissioned to investigate the issue on 16th April 2020. ICCRT published an analysis of clinical outcomes of patients hospitalised in London with COVID-19 on 29th April 2020 [NF/301 - INQ000262569]. This study – based on early and therefore limited data – concluded that black patients had an approximately 2-fold higher risk of death than white patients. However, I believe SAGE only reviewed data on inequity in outcomes by ethnicity from late May, first commenting on this issue on 4th June [NF/302 - INQ000120526].

344. Inequity in clinical outcomes and in relation to interventions (both NPIs and later vaccination) became a higher priority for SAGE from June 2020 onwards. Correlations between both ethnicity and deprivation and local COVID-19 hotspots was noted in the summer of 2020, and an ethnicity subgroup of SAGE was formed in September 2020.

345. ICCRT researchers also contributed to a model-based evaluation of the benefits of the COVID-PROTECT accommodation scheme established for people experiencing homelessness in England in the first wave of the pandemic.

346. Most detailed epidemiological analyses of individual-level COVID-19 data (*e.g.* estimating variants characteristics and vaccine effectiveness) undertaken by SPI-M-O groups (including ICCRT) from June 2020 accounted for differences by ethnicity and deprivation level. However, I am not aware of any SPI-M-O epidemic (*i.e.* mechanistic) models used in the first two years of the pandemic which explicitly represented variation by either deprivation/income group or ethnicity. There are two major reasons for this: (a) model complexity: including, say, 5 broad ethnicity categories and 5 deprivation categories would increase the number of variables being tracked in compartmental models (which nearly all SPI-M-O groups routinely used) 25-fold, making weekly updating of modelling projections computationally unfeasible; (b) data gaps: detailed data on contact patterns stratified by ethnicity and income group were not available (the CoMix survey provided limited summary data, but was not powered to provide the type of detailed data models required). I would note that this area is currently an active research topic for ICCRT (see para 510 section N).

K. Genetic and epidemiological assessment of variants of concern

Introduction

347. The UK was very proactive in initiating and then substantially resourcing genetic sequencing of COVID-19 test samples. For much of the pandemic, the UK was generating more viral sequences than any other country. I attended the 11th March 2020 meeting at the Wellcome Trust in London where the COG-UK consortium [NF/303 - INQ000275160] was initially conceived. ICCRT made significant use of COVID-19 genetic sequencing data throughout the pandemic, with that stream of work being led by my colleague Dr Erik Volz. I would also note that PHE started sequencing of UK COVID-19 samples in January 2020, before the establishment of COG-UK. The UK's investments in sequencing put the country in an excellent position to track and then characterise new COVID-19 variants of concern.

348. It is also admirable that the COG-UK, as conceived by its founder Sharon Peacock and others at the 11th March meeting, was a scientific collaboration spanning dozens of UK researchers, not just a sequencing project.

349. I should note at this point that genetic sequence data for rapidly mutating RNA viruses such as SARS-CoV-2 can, with appropriate analysis, provide valuable insight into the epidemiology of an infection, not just into its evolution. The rate at which mutations accumulate in a virus population is proportional to the overall number of people infected. So-called phylodynamic analysis methods allow sequence data to be analysed to give insight into transmission dynamics. This is particularly valuable when testing is limited or systematic surveillance is lacking.

350. Genetic sequence data can also be analysed to examine whether there is evidence for particular viral variants (more precisely, lineages) having a selective advantage (in a Darwinian sense) over other lineages. The key signature for such an advantage is a new lineage emerging which is then seen in a rapidly increasing proportion of all samples sequenced thereafter. However, in undertaking analyses of these type of signals, it is important to be aware of the possibility of so-called founder effects – namely it may just be random chance that a particular mutation occurred at a time and place (and within a particular sub-population) where transmission rates were high.

351. We published our first phylogenetic analysis of COVID-19 on 15th February 2020 [NF/304 - INQ000236278], which concluded that the virus likely first became established in the human population in China in early December 2019.

352. From the establishment of COG-UK, ICCRT researchers worked closely with the PHE "Genomics Cell" led by Meera Chand, to the extent that Erik Volz and members of his group had part-time secondments to PHE to assist with genetic analyses.

December 2020 – Alpha

353. While the Alpha variant (the B.1.1.7 lineage, also popularly referred to as the "Kent" variant) is the first "variant of concern" which became widely known, it is worth noting that the SARS-CoV-2 virus had exhibited previous mutations with significant epidemiological consequences. The lineage of the virus which dominated the first wave of the pandemic in the UK and most other European countries carried a mutation (614G) not seen in the original Wuhan strain. Retrospective analysis by COG-UK (led by Erik Volz) showed this mutation likely increased viral transmissibility (compared with the Wuhan strain), though not severity.

354. With colleagues in the PHE Genomics Cell and the University of Edinburgh (Andrew Rambaut), ICCRT staff (led by Erik Volz) played a key role in the discovery and characterisation of the Alpha variant. Stimulated by a WHO discussion of the significance of the N501Y mutation seen in some sequences in South Africa a few days earlier, on 7th December 2020, this three-way collaboration identified a worrying signal that the B.1.1.7 lineage in the UK was growing substantially faster than other viral lineages. By 9th December 2020, they had established a statistical correlation at local authority level between the frequency of the B.1.1.7 lineage and the incidence of COVID-19 cases and local growth rates.

355. By 11th December, they had noted the potential for a particular (location 69/70) codon deletion in the spike protein of the B.1.1.7 lineage to be used to obtain more detailed data on the spread of the lineage from routine PCR testing data. This was because the approximately 40% of Pillar 2 PCR testing which went through the three very largescale "Lighthouse" testing labs made use of a commercial PCR test which used three gene targets, including (unusually for COVID-19 tests) the Spike protein. Due to the 69/70 deletion in the Spike protein, the samples from people infected with the B.1.1.7 lineage tested positive on the two non-Spike targets, but negative on Spike (so-called S-gene Target Failure – SGTF). This provided much more data to track the new variant, and data which were substantially more timely (given

sequencing was taking a week or more at that point). SGTF data proved invaluable throughout 2021, as Delta tested positive on Spike while Omicron was again negative.

356. I first became aware of the potential significance of the Alpha variant from a presentation given by the PHE Genomics Cell to NERVTAG on 11th December 2020 [NF/305 - INQ000120390]. Later that day, I alerted Patrick Vallance to what I viewed as a concerning picture. Within ICCRT, we accelerated and expanded our work on the Alpha variant from that time, which involved me and other colleagues supplementing Erik Volz's team. The work involved focussed on using a variety of approaches to quantify the transmission advantage of the Alpha variant compared with early strains.

357. On 14th December, I attended a small group meeting with Patrick Vallance, PHE staff and a number of NERVTAG members to review the evolving evidence for the increased transmissibility of the Alpha variant. I believe (but cannot be certain) that Chris Whitty also attended. Later that day, the Health Secretary made a Commons statement about the new variant and announced that London and large parts of the East and South East of England were moving into Tier 3 restrictions.

358. We continued intensive work with PHE colleagues that week, culminating in a NERVTAG "Bird Table" small group meeting on the morning of December 18th where we and PHE colleagues presented an extended analysis of the transmissibility advantage of the new variant [NF/306 - INQ000120146]. Peter Horby and I then discussed the findings with Chris Whitty, Patrick Vallance and PHE colleagues early that afternoon. Senior ministers were briefed later that day, and the Prime Minister announced new Tier 4 measures covering London, the East and South East in response to the new variant the following day [NF/307 - INQ000086623]. Tier 4 measures were similar in stringency to the second lockdown, and involved closure of non-essential retail and all hospitality and leisure venues.

359. As an aside, I felt that the SPI-M-O consensus statement of 22 December 2020 [NF/308 - INQ000074956] did not accurately reflect scientific knowledge at the time in stating, in reference to the growth advantage of Alpha, "The underlying cause of that faster spread is, as yet, unclear". The statement goes on to speculate about changes in the generation interval; however, the analysis ICCRT had conducted in collaboration with PHE[227] had already demonstrated a shortening of the generation interval could not explain the epidemiological patterns being seen (namely cases of Alpha increasing in local areas where cases of pre-Alpha

COVID-19 were simultaneously decreasing). This is an example (in addition to the acceptance of IFR estimates in February 2020) where SPI-M-O's requirement for analyses to be confirmed by multiple groups arguably delayed consensus being reached. However, in this instance, I do not feel it impacted on the timeliness of policy-making.

360. Together with Wendy Barclay and Peter Horby, I gave evidence to the House of Commons Science and Technology Committee about the new variant on 23rd December 2020.

361. Over the following week we continued refining our analysis of Alpha, including tracking growth rates by region. I was concerned about the high growth rates in hospital admissions being seen in London and the South East (doubling time of 11-14 days), and while there was some indication of slowing in the growth of Pillar 2 daily cases due to the school holidays and escalation of NPIs, I was concerned that re-opening schools in January would risk increasing transmission once more. These concerns were shared by John Edmunds and other SPI-M-O members, and these and other issues were discussed at an ad-hoc SPI-M-O meeting on 30th December. I also discussed my concerns informally with PHE colleagues and Patrick Vallance. My view by 1st January 2021 was that a third lockdown, with schools closed (unlike the second lockdown), was necessary and unavoidable.

362. The third national lockdown in England was announced on 4th January 2021. Unlike the second lockdown, schools remained closed, due to concerns that suppression might not otherwise be achieved given the higher transmissibility of the Alpha variant.

363. I would note that in the two weeks before the third lockdown, areas in Tier 4 were already effectively under measures similar to those introduced on 4th January, given schools were closed for the Christmas holidays. Hence my overall judgement was that government decision-making in response to the Alpha variant (*i.e.* from 14th December on) was reasonably timely and proportionate.

364. However, to re-iterate my comments in section J, it was highly regrettable that infection levels were allowed to reach the levels they did prior to 14th December – something which was a direct consequence of decisions made (and not made) back in September and October 2020. Sixty thousand deaths within 28 days of a COVID-19 test were reported in the UK between 14th December 2020 and 14th March 2021.

365. Our work with PHE on the Alpha variant was summarised in public PHE Technical Briefings issued in December 2020 and January 2021, the first being published on 21st December 2020

[NF/309 - INQ000237371]. A formal preprint [NF/310 - INQ000262561] and journal article followed [NF/311 - INQ000262612]. ICCRT staff then remained highly engaged in research to characterise new variants of concern throughout 2021.

January-November 2021 and Delta

366. We continued working on Alpha throughout January 2021. We conducted a confirmatory analysis indicating that Alpha had higher severity (i.e. the CFR was 30-40% higher) than previous pre-Alpha lineages. This was presented at the 15th January NERVTAG meeting [NF/312 - INQ000120435], at the SPI-M-O meeting of 20th January, and a NERVTAG summary of the evidence on the severity of Alpha [NF/313 - INQ000262599] was considered at the 21st January SAGE meeting [NF/314 - INQ000120586].

367. Early 2021 saw a number of variants of concern become established in different countries: Beta in South Africa, and Gamma in Brazil. Both were a concern as they showed more evidence of immune escape than Alpha, potentially compromising vaccine effectiveness.

368. In February 2021, South Africa stopped using the AstraZeneca vaccine after concerns about vaccine effectiveness against the Beta variant were raised. Given the UK's reliance on the AstraZeneca vaccine, I became concerned about the rise in the proportion of Beta in French COVID-19 viral sequences in early March 2021, which suggested Beta might be outcompeting Alpha, at least in some contexts. I emailed Patrick Vallance and Chris Whitty about this topic on 17th March 2021 [NF/315 - INQ000148851; NF/316 - INQ000148852; NF/317 - INQ000148853; NF/318 - INQ000148854], suggesting that this was one occasion where border measures might be valuable in order to slow the rate of importation of Beta infections from France. I believe there was extensive discussion in government about potentially classifying France as a "red list" country over the next 10 days. The final decision was not to make such a move, given the highly disruptive input it would have on trade, but the testing regime for commercial drivers was intensified. I note that international leisure travel was still banned at that time.

369. The risk posed by Beta was quickly overtaken by that from another variant, Delta (B.1.617.2), which (in retrospect) drove (with other B.1.617 lineages) a very large epidemic in India from March 2021. PHE and its academic partners (including ICCRT) started tracking B.1.617 lineages in detail from early April, with B.1.617.1 being designated a variant under

investigation by 21st April [NF/319 - INQ000120388], and B.1.617.2 (Delta) being classified as a variant of concern on 6th May [NF/320 - INQ000222014].

370. In March 2021, PHE created an advisory group (later named the Variant Technical Group - VTG) to provide expert input into ongoing monitoring of trends in COVID-19 viral evolution and associated risk assessments. VTG met as needed during 2021. In my view, this has been a highly effective group, in part because of its highly multidisciplinary nature.

371. In May 2021 we undertook a detailed analysis (using linked case, death and vaccination data) of the epidemiological characteristics of Delta as it emerged in England. This examined its growth advantage relative to Alpha, differences in the characteristics of infected individuals (e.g. by region, age and ethnicity) between Alpha and Delta cases, and differences in vaccine effectiveness compared with Alpha. We estimated that Delta had a 50-80% transmission advantage (higher R) than Alpha. These analyses were updated weekly for 6 weeks, with updates being fed into the VTG and SPI-M-O each week. Our [NF/017 - INQ000262591] and other SPI-M-O groups' analyses of Delta were considered at the 3rd June 2021 meeting of SAGE [NF/321 - INQ000120623].

372. The emergence of Delta led to a four week delay in Step 4 of the roadmap out of lockdown. This is discussed in more detail in section L below.

373. I note that during 2021, ICCRT staff worked closely with international collaborators on a number of analyses of international data on different variants. Working with Brazilian colleagues, we published analyses of the Gamma variant in March [NF/322 - INQ000273320; NF/323 - INQ000273321], and with Indian colleagues contributed to analyses of Delta in India [NF/324 - INQ000273318; NF/325 - INQ000275161].

374. Over the summer and autumn of 2021 we also worked closely with UKHSA colleagues responsible for vaccine effectiveness (VE) estimates, contributing to the VE working group convened weekly by UKHSA. This work focussed on evaluating the extent to which VE against infection and severe disease was reduced for Delta versus Alpha, and quantifying the waning of vaccine effectiveness over time.

December 2021 – Omicron

375. On 26th November 2021, the B.1.1.529 lineage (Omicron variant) was identified as a variant of concern due to it driving a new epidemic wave in South Africa [NF/326 - INQ000262627]. We initiated a rapid research programme to evaluate the potential effects

of Omicron on vaccine effectiveness, to evaluate the growth rate and transmission advantage of Omicron versus Delta (repeating the type of analyses previously undertaken for comparing Delta and Alpha), and to evaluate the severity of Omicron relative to Delta. Preliminary results regarding vaccine effectiveness and growth advantage were shared from 1st December with SPI-M-O, the VTG and some ad hoc meetings. We shared a more detailed analysis of risk factors for Omicron versus Delta infection (giving insight into levels of vaccine escape) with Patrick Vallance, Chris Whitty, SPI-M-O and JCVI on 6th December. An updated version of the latter analysis was presented in a public report published on 16th December [NF/327 - INQ000262563] and our analysis of relative severity on 23rd December [NF/328 - INQ000212123].

376. I would note that the first quantitative analyses of the relative severity of Omicron compared with Delta were a South African study published on 21st December [NF/329 - INQ000262604] and our own on the 23rd December [NF/328 - INQ000212123]. Our analysis, based on data up to 21st December, indicated an approximately 40% reduction in the risk of hospitalisation, a smaller reduction than estimated for South Africa, in part because the UK population was much more heavily vaccinated than South Africa, and Omicron somewhat compromised vaccine effectiveness.

377. In collaboration with UKHSA and the MRC Biostatistics Unit in Cambridge, we updated and refined our analysis using much more data in early 2022, providing estimates of relative severity for a number of measures of hospitalisation and for death [NF/330 - INQ000262572]. This analysis indicated an overall 44% reduction in the risk of hospital attendance, a 59% reduction in the risk of hospital admission, and a 69% reduction in the risk of death (comparing Omicron cases with Delta cases). It also provided age-stratified estimates of relative severity, showing that there was little or no difference in severity (between Omicron and Delta) for individuals less than 20 years of age.

378. We continued to participate in the UKHSA VTG after February 2022 and to contribute to a number of analyses.

L. Vaccination modelling

Note: with the agreement of the Inquiry, factual descriptions of activities undertaken by ICCRT provided in this section were written with the assistance of Drs Anne Cori, Lilith Whittles, Natsuko Imai and Marc Baguelin. All statements of opinion are entirely my own.

October-December 2020

379. We started developing models incorporating vaccination in July 2020, initially to inform WHO and global planning, and later to support vaccination planning in specific countries, including the UK. In this section, I will focus on UK modelling, but further details of the modelling for international organisations can be provided on request.

380. We started UK-focused vaccination modelling in November 2020 following the first promising interim results from vaccine clinical trials. We expanded the SirCovid model to include vaccination. The representation of vaccination in the model was refined throughout 2021-22 to account for delays from vaccination to the onset of effectiveness, waning of vaccine-induced protection over time, and 3rd "booster" doses. We assumed vaccination induced four types of protection: 1) reduced risk of infection, 2) reduced risk of severe disease given infection, 3) reduced risk of death given severe disease and 4) reduced infectivity (i.e., onward transmissibility).

381. Commissioned by SPI-M-O, in November 2020 we examined multiple scenarios considering both vaccination and NPIs. This work was considered within a sub-group of SPI-M-O dedicated specifically to vaccination, which included three modelling groups (Imperial, LSHTM, and Warwick). Scenarios considered were iterated and refined over time, with reports shared with SPI-M-O and results from the three groups discussed and compared roughly weekly or fortnightly.

382. In November and December 2020, we shared our first results on the potential impact of a UK COVID-19 vaccination campaign [NF/331 - INQ000114467]. This followed a SPI-M-O commission on vaccination and NPIs from 13th November 2020, which then led to a small subgroup of SPI-M-O focusing on vaccination which first met on 20th November 2020. As evidence on real world vaccine effectiveness and potential vaccine uptake was limited at the time, we considered a wide range of vaccination scenarios. We assumed a mix of Pfizer and AstraZeneca vaccines, varying vaccine uptake and efficacy, and the nature of the vaccine-derived protection. We also considered several scenarios regarding future levels of transmission depending on when and how NPIs (such as the Tier system) would be changed. Despite promising vaccine trial results, we concluded that vaccination would not immediately allow lifting of all restrictions without inducing a new wave of infections and hospitalisations. In all scenarios considered, high vaccine uptake across all age groups would be required to

sufficiently protect the population and some NPIs would need to remain in place through 2021 to prevent a significant increase in the number of infections, hospitalisations, and deaths. Our results (together with similar analyses from the Warwick group) were discussed at SAGE on 10th December 2020 [NF/331 - INQ000114467].

2021 and the Roadmap out of lockdown

383. Throughout January 2021, we (and the LSHTM and Warwick groups) produced further scenario analyses to include more detailed information about the anticipated vaccine rollout speed. These were discussed at SPI-M-O on 6th and 20th January and at SAGE on 14th January 2021 [NF/332 - INQ000075533; NF/333 - INQ000262608]. Vaccine effectiveness assumptions were also refined as real-world data became available against the newly emerged Alpha variant and previously circulating lineages. We expanded the SirCovid model to include two variants to explicitly model Alpha. Our results demonstrated the need for some level of NPIs to remain in place until the summer of 2021 to avoid a third wave of hospitalisations even larger than the second wave, and highlighted the benefits of accelerating vaccination rollout, particularly considering the growing Alpha epidemic.

384. In January 2021, SPI-M-O commissioned the three groups undertaking vaccination modelling to model the projected impact of stepwise approaches to relaxing NPIs at various predefined paces, from "very fast" to "gradual". The SPI-M-O subgroup on vaccination met very frequently in this period up to the announcement of the roadmap out of lockdown policy (6 meetings between 3rd and 18th February 2021, plus a number of email exchanges), to refine scenarios and discuss results from the three groups. We modelled multiple scenarios reflecting these hypothetical NPI lifting plans and demonstrated that it would take several months to achieve high levels of protection in the population [NF/006 - INQ000262592; NF/334 - INQ000063695; NF/335 - INQ000063699]. This was based on the anticipated vaccine roll out speed, vaccine eligibility criteria at the time, predicted vaccine hesitancy levels, the need for two vaccine doses to be fully protected, and the delay between vaccination and developing a level of immunity. For example, we projected that it would take until July 2021 for half of the population to have vaccine-induced protection against mild disease (see Figure below). We showed that to avoid a large epidemic rebound, the timing of NPI lifting should be aligned with the gradual increase in vaccine-induced protection in the population.

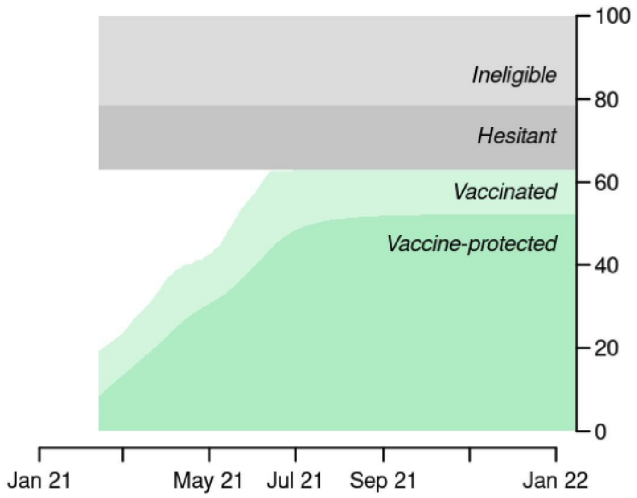

Figure: Projected proportion of the population in England protected against severe disease through vaccination over time (dark green shading) and vaccinated (having received one dose) over time (light green shading). The grey shaded areas show the proportion of the population ineligible for vaccination (i.e. <18 years, light grey) and those who are vaccine hesitant and not taking the vaccine (dark grey). Figure reproduced from [NF/335 - INQ000063699].

385. Thus, we suggested that NPI lifting should be implemented in a gradual stepwise manner, re-evaluated after each step to monitor transmission (with sufficient time between steps to allow for proper evaluation), and revisited if needed – in particular should vaccination roll-out be slower than anticipated. This work was discussed at SAGE on 4th [NF/336 - INQ000086985], 11th [NF/006 - INQ000262592] and 18th February 2021 [NF/335 - INQ000063699].

386. On 22 February 2021, the UK government announced the roadmap out of lockdown [NF/337 - INQ000237613], a plan to lift restrictions in 4 non-reversible steps (step 1, schools reopened, outdoor gatherings of no more than six people allowed; step 2, non-essential retail opened; step 3, outdoor gatherings up to 30 people allowed, indoor gatherings up to 6 people allowed; step 4, no limits on social contact). Tentative dates were announced for each step (step 1: 8 March, step 2: 12 April, step 3: 17 May, step 4: 21 June), with the caveat that these dates could be delayed, as each step would only be undertaken if four conditions were met: 1) vaccine roll-out is as anticipated, 2) vaccine effectiveness against severe disease and death is sufficiently high, 3) infection rates and hospitalisations remain low and 4) the risk assessment is not modified by new variants of concern (VOCs).

387. I felt the roadmap represented a clear, evidence-based approach with well-defined long-term strategic objectives – namely to relax restrictions as rapidly as possible while minimising the risk of an additional wave of transmission which could risk overwhelming healthcare

capacity or require reimposition of restrictions. I would note that it prioritised social and economic reopening over absolute minimisation of COVID-related mortality or hospitalisations, but I felt that was a choice for policy-makers to make.

388. From March 2021, SPI-M-O commissioned a series of reports (from our group, LSHTM and Warwick) to monitor and evaluate the epidemic at each step of the roadmap [NF/338 - INQ000063702; NF/339 - INQ000262609; NF/340 - INQ000063701]. The aim was to estimate transmission levels after each NPI lifting step, compare these with the assumptions in our modelling to support the design of the roadmap, and produce updated epidemic projections. Projections included updated vaccine roll-out information incorporating the latest emerging evidence on vaccine effectiveness as well as assumed social mixing levels. We also considered the hypothetical emergence of a variant of concern and potential implications for the roadmap. This series of analyses was presented at multiple SPI-M-O subgroup meetings (12th and 30th March, 4th and 17th May, again with many emails exchanged in addition), and discussed at SAGE on 31st March 2021 [NF/338 - INQ000063702; NF/340 - INQ000063701].

389. From mid-May, following detection of the Delta variant in the UK and with Delta likely to become the dominant variant (see section K), we explored the potential impact of Delta on the roadmap [NF/341 - INQ000262610]. This was discussed at SPI-M-O subgroup meetings on 8th June and 6th July 2021. Despite large uncertainties around Delta's transmissibility, severity, and immune escape properties, initial analyses demonstrated that lifting the remaining NPIs as planned on 21 June 2021 (step 4) could lead to a third wave of hospitalisations at least as large as the previous winter wave (this was discussed at SPI-M-O on 19th May 2021). In response to a commission from SPI-M-O, from early June we (and other groups) projected the impact of delaying the 4th step of the roadmap, between 2 weeks and six months [NF/342 - INQ000092672; NF/343 - INQ000120630; NF/344 - INQ000262611]. Despite large uncertainties in the magnitude and timing of the coming epidemic wave, we showed that a delayed step 4 would enable more vaccine doses to be distributed, which in turn would substantially reduce hospitalisations and deaths. We also highlighted that a delay in step 4 of the roadmap would provide time to better quantify Delta's characteristics and refine epidemic projections. This work was discussed at SAGE on 9th June 2021 [NF/341 - INQ000262610], and later on 7th July 2021 [NF/343 - INQ000262611][5].

390. On 14th June 2021, the government announced step 4 would be delayed by four weeks until 19th of July, and that the vaccination campaign would be accelerated, with the interval

between doses reduced from 12- to 8-weeks [NF/345 - INQ000086733]. We retrospectively estimated that the delay in step 4 of the roadmap reduced peak hospital admissions by approximately three-fold [NF/346 - INQ000262600].

391. The arrival of Delta in the UK began a five-month period of high sustained case incidence. Case numbers averaged over 25,000 per day from the beginning of July 2021 to the beginning of December, at which time Omicron started to replace Delta.

392. Between mid-May 2021 and mid-July, coincident with the relaxation of NPIs and the replacement of Alpha with Delta, UK case numbers rose from under 2000 per day to a peak of over 60000/day. Both our and SPI-M-O medium-term projections at the time suggested that the UK would see over 100,000 cases per day within one or two weeks, leading me to make the ill-advised statement (in a BBC Andrew Marr show interview) that it was "almost inevitable" that the UK would see 100,000 cases per day. Case numbers then fell continuously for the two weeks after I gave my interview, before stabilising at an average of about 25,000 per day from the start of August.

393. Later SPI-M-O analyses of those few weeks suggested that the Euros football tournament may have led to a significant but temporary increase in population contact rates in June and early July 2021. This was also the period when the NHS COVID-19 app had highest use, with many media stories about the resulting "pingdemic", so it is possible that the very large number of app notifications to self-isolate may have significantly slowed transmission in July 2020. However, SPI-M-O groups (including ICCRT) still don't have a definitive explanation for the rapid rise in infection levels between May and July and then their later plateauing at high levels for 5 months – a trend seen both in case data and the ONS and REACT infection surveys. The long plateau period was not a pattern seen in other European countries, and implies an almost exact balancing of immunity levels and contact rates in the population (to give R=1) – not something we (or any SPI-M-O group) predicted in advance. This experience further reinforced our collective understanding of the limits of epidemic forecasting and taught me a personal lesson in being more cautious in public statements about future epidemic trends.

394. Throughout 2021, vaccine effectiveness assumptions used in the SirCovid model were constantly updated. Most values were informed by UK-based studies published by UKHSA throughout the pandemic. These were supplemented by international vaccine effectiveness

studies (e.g. Israel) which helped to inform the level of uncertainty around effectiveness assumptions which were explored in sensitivity analyses and a wide range of scenarios.

395. On 22nd September 2021, we were asked by SPI-M-O to examine the impact of the full reopening in England during the coming winter, even though no formal "step 5" had been planned. The purpose of this request was to assess potential contingency planning for the winter. In response, we submitted a report to SPI-M-O on 13th October 2021 that summarized potential trajectories for the COVID-19 epidemic until March 2022, based on recent data and assumptions regarding changes in contact rates throughout the winter, vaccine effectiveness and coverage, cross-protection between variants, and the waning of infection- and vaccine-induced protection. This modelling, together with parallel work by LSHTM and Warwick was summarised in a SPI-M-O paper [NF/347 - INQ000087023] considered by the 14th October 2021 SAGE meeting [NF/348 - INQ000120650].

396. On 27th November, following the emergence of the Omicron variant, we received a request for some emergency work by SPI-M-O for the following Monday 29th SAGE meeting. On 29th November I emailed a report to Patrick Vallance and Chris Whitty which projected the potential drop in vaccine efficacy due to the possible immunity escaping property of Omicron [NF/349 - INQ000148855; NF/350 - INQ000148856; NF/351 - INQ000148857; NF/352 - INQ000148858; NF/353 - INQ000148859; NF/354 - INQ000148860; NF/355 - INQ000148861]. These vaccine efficacy projections were later discussed with JCVI.

397. On 6th December 2021, I emailed the results of preliminary modelling of potential Omicron epidemic scenarios to Patrick Vallance, Chris Whitty, UKHSA and the SPI-M-O and SAGE secretariats [NF/356 - INQ000148862; NF/357 - INQ000148863]. This report combined data on Omicron growth rates, neutralizing antibody titres (NATs measuring the degree of immune escape), and risk factors for Omicron versus Delta infection to generate a large number of epidemic scenarios for Omicron. One of the main aims of our analysis was to examine differences in outcomes according to the extent to which the growth advantage of Omicron was due to vaccine escape versus increased transmissibility. This did not include any modelling of reintroduction of NPIs or extension of vaccination programmes (*e.g.* additional boosters, or vaccination of 5-11 year-olds.)

398. All our scenario modelling for Omicron assumed Omicron and Delta had the same severity. The reports and presentations presenting the results noted this was a key

uncertainty, and my email stated that model projections of hospitalisations and deaths for Omicron could be scaled proportionally if the relative severity differed from Delta [NF/356 - INQ000148862; NF/357 - INQ000148863]. We had separately started an analysis of UK surveillance data to assess the relative severity of Omicron at that time (see section K above).

399. Based on our epidemiological analyses (see section K above), we explored scenarios where the exponential growth rate of Omicron infections varied between 0.2-0.45/day and that Omicron could lead to a 2- to 4-fold reduction in NATs compared with Delta. Our modelling suggested that to reproduce epidemic growth rates being seen at that time, a four-fold decrease in NATs would have required at least a 25% increase in Omicron's transmissibility compared with Delta. Most scenarios consistent with currently estimated growth rates for Omicron were projected to lead to a wave of infection with very high peak daily infections and case numbers. Under the assumption of Omicron and Delta having the same severity, this led to high hospital demand and at least 28,000 additional deaths by June 2022 compared with the "no Omicron" scenario. However, somewhat more moderate scenarios were plausible if the NAT reduction was closer to 2-fold than 4-fold, and if the transmission advantage of Omicron was principally due to greater intrinsic transmissibility.

400. On 7th December 2021, we joined a SPI-M-O small group meeting to discuss the scenario modelling papers from our group, LSHTM and Warwick to produce consensus SPI-M-O statements [NF/358 - INQ000218240; NF/359 - INQ000074963][4] on potential impacts of the Omicron wave. There was a broad consistency in the modelling results from the three academic groups. Additional modelling of Omicron intervention scenarios for SAGE was commissioned at that meeting, but due to the demands of our other analytical work on Omicron (see section K) and the extent of changes we had to make to the SirCovid code to model these (notably being able to model Omicron and Delta having different generation intervals), we were not able to submit a finalised report by the deadline for this request (though we did present some preliminary work at the 15th December SPI-M-O meeting in PowerPoint format). Hence only modelling from LSHTM and Warwick were included in the documentation for the December 16th 2021 SAGE meeting [NF/360 - INQ000120658].

401. I was not involved in policy-focussed discussions around Omicron, but felt that announcement of the move to Plan B made on 8th December [NF/361 - INQ000086632] was a proportionate response to the threat posed by the new variant, given the absence of reliable data on relative severity at the point the decision needed to be made.

402. While we continued to provide Medium Term Projections every other week to UKHSA throughout 2022, we have undertaken no additional prospective modelling of intervention options for SPI-M-O since December 2021.

M. The UK science-policy interface during infectious disease crises

Introduction

403. I preface my remarks by commenting that I think that, in most respects, SAGE-COVID was remarkably effective under sometimes challenging circumstances. It was a real plus that it answered to both the GCSA and CMO; in 2009, each official had run their own advisory committee, something which did not improve the coherence of science advice into government in that pandemic.

404. More generally, Patrick Vallance and Chris Whitty (together with the major funders such as UKRI and NIHR) were instrumental in ensuring the UK punched well above its weight in research on COVID-19, particularly in the development and evaluation of treatments and vaccines.

405. However, some aspects of what SAGE ended up doing were less than optimal, given its ad hoc nature. It ended up almost acting as a standing epidemiological assessment group and shaped policy to a greater extent than was optimal – causing tensions between it and some other parts of government at times. The move to publishing SAGE meeting summaries, while admirable from a transparency perspective, at times exacerbated this tension – most notably in September and October 2020. Furthermore, SAGE developed into a substantial government structure (with multiple sub-groups and task-and-finish groups) in its own right as the pandemic progressed, with the implicit role of providing the definitive scientific input into government. I don't believe this was optimal, given SAGE's position outside normal government departmental structures and its ad hoc and intended transient nature.

406. A key challenge for SAGE in the first year of the pandemic was that many of its deliberations lacked clearly stated overarching policy objectives and government "red lines". For instance, SAGE was never explicitly asked to evaluate what policies would lead to minimum use of economically and socially disruptive NPIs while avoiding hospital admissions

or occupancy breaching a specific threshold. Or – and perhaps this would have been a more appropriate use of SAGE – to review a range of strategic policy options drawn up by government and to provide scientific challenge as to their suitability in meeting the stated policy goals. More generally, there was relatively little systematic communication of policy deliberations or goals down from COBR to SAGE – though some information percolated through informal discussions. Hence a lot (if not the majority) of SAGE discussions focussed on reviewing the immediate epidemiological situation (something it would have better for SAGE not to have really needed to do) and individual intervention options (e.g. masks, bubbles, mass testing).

407. More generally, I think the pandemic highlighted a number of key limitations in how the UK government plans for and responds to contingencies. I detail these limitations below and suggest some possible actions to address them. I also note that I understand that much (but perhaps not all) of what I recommend in relation to UKHSA already forms part of current plans; the notable exception might be the creation of a substantive, rigorous and quantitative policy analysis and assessment capability.

The UK model for planning for and responding to civil contingencies

408. While the COBR mechanism to respond to civil contingencies has been in existence since the 1970s, SAGE as a formal structure for gathering real-time scientific advice to inform COBR decision-making emerged out of the experience of the 2001 Foot-and-Mouth Disease (FMD) epidemic in UK livestock. During that epidemic, the then GCSA, David King, assembled an ad-hoc scientific advisory group (I was a member) to improve situational awareness and policy-relevant scientific advice going into government in the management of that crisis.

409. I will not review the history of the FMD epidemic here, though I believe it is of fundamental relevance to module 1 of the Inquiry. However, it is of note that the 2001 advisory group often considered operational aspects of policy in some detail.

410. It is also notable that reviewing epidemiological analysis and modelling made up a substantial part of SAGE activities during the 2009 Influenza pandemic, the 2014 West African Ebola epidemic as well as the COVID-19 pandemic. In this, SAGE followed the pattern set by David King's FMD scientific advisory group.

411. It was only when the ad-hoc structures used in 2001 evolved into the formal SAGE mechanism that a clear division was made between scientific advice (the responsibility of

SAGE) and policy-making, operational planning and implementation (the responsibility of COBR, lead departments and technical agencies). While I understand the rationale for it, trying to maintain this division of responsibilities has been a tension ever since.

412. I think this issue is symptomatic of broader issues in the UK's governmental mechanisms for responding to certain types of crises. So rather than just focussing on SAGE, I feel the UK needs to rethink its response to crises overall, and then plan advisory structures in that broader context.

413. I will therefore discuss the broader context of how the UK's approach to risk assessment and crisis response has evolved in the last 20+ years and the weaknesses revealed by the pandemic.

414. I am not an expert in civil contingency planning, decision analysis or risk assessment. However, I have sat on SAGE and its predecessor during multiple different infectious disease crises over more than 20 years, and have had significant interaction with civil servants in multiple government departments on risk assessment, preparedness and response over that time.

415. The creation of the Cabinet Office Civil Contingencies Secretariat (CCS) in the aftermath of FMD and 9/11 presaged a more systematic approach to crisis preparedness in the UK government.

416. In the years after its creation, the CCS spearheaded a review of major risks and the creation of the National Risk Register (NRR), an assessment of the major potentially significant risks facing the UK. This was initially a classified assessment, but from 2008 a public facing version was published.

417. I was involved in discussions around the methods used to quantify risks in the NRR on a number of occasions since 2001, and advised specifically on a number of the key risks identified.

418. The NRR quantifies risk within a two-dimensional "likelihood"/"impact" matrix. Impacts are assessed in a cross-sectoral manner, spanning mortality, the economy, essential services and population displacement.

419. Risks are assessed and located on the matrix on a RWC basis. This is reasonable but perhaps not optimal in terms of planning policy responses, as I'll discuss in more detail below.

420. Risks identified in the NRR are the topic of cross-government risk mitigation planning, typically with a lead department coordinating planning for each risk. Mitigation is focussed on preparedness and response planning to reduce impact, though sometimes reducing likelihood is also a focus.

421. In general, I think the UK has been ahead of many countries in this quantitative and relatively rigorous approach to risk assessment and management.

422. For many (if not all) of the risks identified, a "playbook" of policy responses has been formulated. There has also been an emphasis – accelerated by the COVID-19 pandemic – on better understanding of cross-sector network ("domino") effects, and on real-time data capture to inform situational awareness (e.g. real-time dashboards presenting key indicators).

Lessons from COVID-19 for policy planning

423. However, I think the COVID-19 pandemic (and before that, the 2009 Influenza pandemic) have highlighted some limitations of the current approach to contingency/extreme event planning and response in the UK.

424. I believe that the focus on RWC scenarios is overly restrictive, and risks leading to tunnel-vision – both in only planning for "very bad" scenarios, but also in perhaps giving officials and policy-makers the sense that reality is unlikely to be as bad as the RWC.

425. A related challenge is that for certain risks on the NRR (particularly pandemics, but also some other), it is not at all straightforward to assess the (potential) impact of an event immediately. Indeed, it is often not obvious that the relevant event *has* started; *e.g.* that a potentially concerning infectious disease outbreak in another country really poses a risk to the UK.

426. Focussing specifically on epidemic/pandemic risks, as an outbreak unfolds, there will always be a need to iteratively reassess the costs and benefits of policy actions or *inaction* in a context of likely high levels of uncertainty.

427. While the judgement of what constitutes an "appropriate" policy response (as compared with over-reaction or under-reaction) will always be a political one, I think the government can do much better than it has in the past in prospectively mapping out potential epidemic scenarios and policy responses for each.

428. This will require moving beyond viewing risks through the lens of the RWC, but instead assessing the range of impacts a novel infectious disease threat (which might or might not cause a pandemic) could generate. This would probably best be done as a set of scenarios, spanning expected ranges of transmissibility, severity, but also more categorical variables (e.g. mode of transmission, risk profile across age groups). Then for each scenario, an evidence-based menu of policy responses can be generated, spanning a range of risk appetites from the precautionary "escalate first, de-escalate if needed" approach to the "never cry wolf" end of the spectrum.

429. Critical to such an approach is a detailed consideration of how to balance the scale of potential impacts of both the disease (i.e. deaths) and the policy response (i.e. economic and social disruption) and the uncertainty in the assessment of those impacts existing at a point in time.

430. Both SAGE and UKHSA use a sensible semi-quantitative system for evaluating the confidence they have in particular pieces of evidence (e.g. an estimate of the IFR), particularly within formal risk assessments. This involves a categorical scale which spans the range from very low confidence (meaning highly uncertain or depending on a single evidence source) to very high confidence (multiple convergent streams of high quality evidence).

431. In my view, much more thought needs to be given to how those assessments of uncertainty/confidence influence decision-making, especially in the context where the "central" estimate points towards a very high impact event. The default policy response to the high levels of uncertainty we saw early in the COVID-19 pandemic was to prioritise collecting more data, to update RWC scenarios, but in other respects to wait and see.

432. Much of the work I propose above can be done now. Research should be commissioned to systematically compare the health and economic costs of waiting too long to act against those of over-reacting early and then needing to de-escalate. This can draw upon the enormous amount of data which has been collected on the health, social and economic impacts of the pandemic, and on the effectiveness and socioeconomic impacts of NPIs.

433. A key factor for such research to consider is the *time horizon* over which policy and disease impacts are assessed. In my view, a critical failing in the UK policy response to COVID-19 in 2020 – particularly between May and December – was the overly short time horizon over which the impacts of NPIs on the economy (and the epidemic, to a lesser extent) were being

evaluated. Adopting a strategy where NPIs were only intensified when the NHS capacity (locally or nationally) was at risk of being overwhelmed was intended to minimise the imposition of economically (and socially/politically) costly measures. However, delaying intensification of NPIs just led to more hospitalisations and deaths without any economic savings, given that measures did eventually need to be escalated, and then kept in place for at least as long (if not longer) than they would have been had action occurred earlier.

434. Of course, being willing to make decisions on the basis of projected costs and benefits evaluated over a, say, 6-12 month time horizon requires policy-makers to trust (and be able to explain) the modelling underpinning those assessments. By comparison, decisions made late can be justified by pointing to recent numbers of hospitalisations and deaths – which was generally what ministers, the GCSA and CMO emphasised when announcing and justifying intensification of NPIs during the pandemic. This speaks to a more general and complex issue – namely that, the longer the time horizon used, the higher will be the uncertainty in evaluating costs and benefits.

435. Nevertheless, even accepting that politicians will vary in their appetite to accept modelled projections versus recent data trends, I feel the formal analysis and presentation of the potential costs and benefits of policy options over a range of time horizons would be of value in clarifying the trade-offs associated with different approaches to interventions.

436. A related critical issue is what the strategic goals of policy (and policy-makers) are (see para 350 also). This was less clear than it might have been for much of 2020. Reading between the lines of government documents introducing the COVID-19 alert levels system in May 2020 [NF/273 - INQ000086784], it would appear that the government wanted to minimise the duration and socioeconomic impact of NPIs while preventing NHS capacity from being overwhelmed. As I have explained above, I don't feel the policies adopted from May to December 2020 came close to achieving that goal.

437. In my view, one of the lessons to learn from that experience is to have technical science advice inform operational policy formulation to a greater extent than happened in 2020. SAGE and SPI-M-O were not prospectively asked for input into or assessment of a number of key policy initiatives, most notably the March 2020 COVID-19 Action plan, the May 2020 Alert level system, the October 2020 local tier system (including the precise criteria for escalating local tiers), and border measures throughout the pandemic. This led to flaws in all of these

policy measures, in my view. While I am not arguing that all policy initiative should have been "approved" by SAGE, a rapid scientific and epidemiological assessment of the likely effects of all of them could have been undertaken before they were introduced. Optimally, such assessments would have been undertaken by UKHSA rather than SAGE directly, but it would be important for them to have been transparent (i.e. published).

Governance of science inputs into government in epidemics

438. Even exhaustive scenario planning will not precisely anticipate the exact characteristics of the next infectious disease threat the UK faces. There will always be a need for real-time and continuous assessment of the actual threat and similarly agile updating of policy playbooks (the "menu" of options). In my view, doing this well is not aided by the current "Chinese wall" between SAGE (scientific advice) and COBR/government (policy planning). Nor can high level committees such as SAGE and COBR necessarily give the level of detailed consideration that are needed to such assessments.

439. I think that technical assessments of risk (including current uncertainty) and of the potential impacts of policy options in mitigating that risk should be the proper responsibility of the relevant lead government technical agency. In the case of infectious disease risks, this is UKHSA. That did not happen in the first 3-6 months of the COVID-19 pandemic in the UK, where PHE did not have the capability to provide the scale and depth of analysis required and SAGE therefore largely took on that role.

440. Placing primary responsibility for scientific inputs (including policy assessment) into government during a pandemic onto UKHSA mirrors systems in place in many northern European countries (Netherlands, Norway, Denmark, Finland and Sweden).

441. As I detail in section N, I strongly feel that UKHSA (and for other NRR risks, other technical agencies) should be proactive in forming long-term partnerships with academic and commercial partners and draw upon the additional capacity those partnerships provide in crisis situations. It will also be key for UKHSA and NHSE to work seamlessly together, particularly in surge capacity planning (and in determining, at an early stage, what the limits are on surge capacity).

442. In the case of a potential pandemic threat, impacts of both the event and policy responses will go well beyond just the health sector, so it would be beneficial for a cross-governmental taskforce with executive authority (determined by COBR) and significant analytical and

technical capacity to be convened at an early stage. Akin to "Gold command" used for terrorist threats and other risks on the NRR, but configured for a potentially much longer duration event. I did not have good visibility of these types of structures within government during the pandemic, but from what I did glean, they were weak in the first 3-4 months and strengthened notably thereafter.

A revised role for SAGE

443. With such a system in place, the need for SAGE should be substantially reduced, but not eliminated. The majority of scientific data which went into SAGE (sometimes for "approval") during the COVID-19 pandemic would better be fed into UKHSA and/or the operational "Gold command" structures. This would leave SAGE with more of a challenge and feedback function; identifying weaknesses and gaps in the evidence base being generated by UKHSA and others, challenging assumptions and even paradigms, and highlighting research priorities or other evidence needs.

444. In doing so, thought should be given to the potential value of SAGE taking on some responsibility for "red teaming" activities, particularly in relation to gap analysis ("what have we missed?"). This role could be adopted by a parallel group, but this might pose challenges in terms of establishing hierarchy and clear lines of responsibility. If red-teaming is to be part of SAGE's function, then its mode of operation and chairing may need to be reassessed. I also note that red-teaming is perhaps difficult if SAGE retains the role it had in the COVID-19 pandemic as the definitive source of scientific evidence for government. It is more compatible with a challenge and feedback role.

445. Red-teaming could also (or alternatively) play a greater role within the more technical advisory structures operated by UKHSA. However, to be fair, my experience of a number of UKHSA technical advisory groups in the last two years is that they have already encouraged a high level of constructive challenge.

446. The disciplinary scope of SAGE should be reconsidered; in particular, I see no reason why the remit of SAGE should not extend to economics and the social sciences more generally. In choosing economists to participate in SAGE, thought should be given to ensuring linkage with Treasury and Bank of England advisory structures.

447. While the role of SAGE is to provide high-level scientific advice to COBR, it is not clear to me that it is always optimal for that advice to be filtered solely through the GCSA and CMO.

This is not an implied criticism of either Patrick Vallance or Chris Whitty, but I see risks associated with having just one or two individuals (plus a civil service secretariat) be solely responsible for communicating complex, policy-sensitive syntheses of scientific knowledge to ministers and senior civil servants.

448. In that context, while the primary distinction between advising on and advocating for a policy should be retained (i.e. "advisers advise, ministers decide"), removing the somewhat arbitrary distinction between strategic/scientific advice and operational/technical advice would be beneficial. This would also allow something of a lowering of the Chinese wall between SAGE and COBR. Occasional joint meetings of participants in both groups might be considered, where SAGE as a committee can discuss scientific assessments with ministers and be questioned on them.

449. That said, I do not have a perfect model of SAGE in mind. However, I think there would be value in evaluating the range of approaches taken across European countries. Ideas to consider include:

a. Having SAGE chaired by one or two independent scientists, while retaining the role of the GCSA (and CMO) in determining the priority list of topics SAGE is tasked with addressing. This may be particularly valuable for red-teaming activities. It would also align SAGE with the practice of statutory committees such as JCVI, ACDP and NERVTAG, all of which have independent chairs.

b. Offering more opportunity for extended scientific discussion within SAGE. Compared with 2009 and 2014, SAGE-COVID was much larger, and run in a considerably more formal manner, particularly as time went on. Much of its activity was reviewing and approving documents, with limited associated discussion. This made brain-storming/horizon-scanning discussions more difficult, which I believe is the main reason why Patrick Vallance organised informal small group discussions on various occasions in 2020-21.

c. Reviewing the role of SAGE participants. "Participant" and "member" perhaps imply different things, especially in relation to collective responsibility and the level of commitment required. In addition, while perhaps unavoidable to a degree, the extent to which scientists on SAGE are providers versus reviewers of scientific evidence should be considered (see section N).

d. Introducing detailed minuting of plenary SAGE meetings, in addition to the summaries currently produced. It is a policy decision as to whether these are published at the time, but overall I believe detailed minuting will aid transparency, give more insight into areas where there is less consensus between SAGE participants, and allow Inquiries such as yours to better understand deliberations retrospectively.

450. I think the expansion of SAGE to include multiple sub-groups and task-and-finish groups should be resisted in future long-duration crises. SAGE became a significant "operationalised" government structure in its own right during the COVID-19 pandemic, almost being viewed as the normative source of scientific input into government. I am aware that Patrick Vallance tried to resist this trend, and that an increasing number of those functions were moved to UKHSA in late 2021. However, my view is that UKHSA – supported by a network of university partners – represents a more sustainable and appropriate source of scientific input into government for an infectious disease threat of potentially multi-year duration. SAGE requests for additional evidence would be better addressed by UKHSA than by SAGE setting up parallel structures itself.

451. This is not to say that I think that government will not benefit from the input of university scientists just as much in the next pandemic as occurred during the COVID-19 pandemic. Rather, I think the vast majority of engagement between government and the academic community should be via UKHSA, with SAGE having the lighter-weight, more challenge-oriented function detailed above.

452. UKHSA will need to be adequately resourced to deliver this function. In particular, I don't think resourcing of pandemic preparedness and decision analysis/policy evaluation is currently adequate. In addition, while there have been significant organisational improvements with the transition from PHE to UKHSA, some unhelpful aspects of past culture remain. Among these is the contain/delay/mitigate "playbook" for handling the earliest stages of a new infectious disease threat which relies too much on measures for while the evidence base is weak.

Transparency

453. While I understand the pressures which led to SAGE participants being named and SAGE documentation being (retrospectively) published, it did lead to the somewhat anomalous situation that SAGE was a transparency outlier within the government apparatus (headed by

COBR) which advised on and developed policy. No such transparency was provided on the deliberations which led, for instance, to the local tier system in October 2020.

454. So in rethinking crisis response in the UK, I think there needs to be a greater emphasis on transparency in decision-making more generally, accepting that there will always be some political and national security considerations. To be more specific, I believe that even if political/policy discussions remain private, the evidence base going into those policy discussions should be made public, wherever possible – accepting that in the heat of the moment, that may need to be retrospective. This is perhaps even more critical if government technical agencies, such as UKHSA, are taking the primary role in making such assessments – as I advocate above. In that way, the deliberations of advisory groups such as SAGE, NERVTAG and JCVI would cease to be the outlier in openly publishing evidence.

455. In that context, thought should be given to how government can provide better pastoral and moral support to academic scientists involved in supporting the UK response to an infectious disease threat. Many participants in SAGE and SPI-M-O (and members of JCVI and NERVTAG) – who were already working under stressful levels of time pressure – were subject to extensive abuse via social media, emails and sometimes mainstream media. From late autumn 2020, GoS provided some pastoral support for SAGE and sub-group participants, but this could be more systematically planned for future events. I would also note that it was disappointing that ministers and senior officials did relatively little to defend scientists from abuse and criticism in public, while praising their contributions in private.

456. From March 2020 onwards, I thought the UK did an excellent job in publishing detailed surveillance data on the pandemic. The UK Government COVID-19 dashboard was clear, timely and comprehensive. Also, since late 2020, PHE/UKHSA has published timely risk assessments and technical briefings on COVID-19 variants and other infectious disease threats.

457. However, my recommendation for greater transparency is not just in relation to surveillance data and risk assessments, but for technical assessments of different policy options (or at least, the adopted policies) to be made public. Policies should be backed by evidence and a realistic and preferably quantitative assessment of what they might achieve. While there were some instances where this happened during the UK COVID-19 pandemic

(particularly from 2021 on), there are some notable counterexamples between May and November 2020 (see section J).

Specific policy recommendations

458. *Border measures*: I think research needs to be commissioned to critically evaluate what the measures adopted at various stages of the pandemic achieved. This should be followed by development of a more evidenced-based strategy for border surveillance and controls for future infectious disease threats. We should never again use the term "containment" to describe a set of measures that are palpably unfit to achieve the goal implied by that term. At the end of the day, some policies with a weak evidence base may still be adopted, but this limitation should be transparent to everyone.

459. *Surveillance and testing*: government should not assume that sustained person to person transmission of a new infectious disease threat isn't occurring in the UK just because it hasn't been detected. In the absence of data to say otherwise, the default assumption should be that it *is* occurring. Surveillance should be rapidly put in place to detect cases unlinked to travel with a certain level of sensitivity, and public statements interpreting the detection (or not) of "sporadic" cases should be couched with reference to the sensitivity of the surveillance system in place. In my view, NHS hospital testing capacity should be maintained at a level which allows rapid (~7 days) initiation of comprehensive hospital-based surveillance for any new infectious disease threat.

460. *Policy red lines*: From my perspective, it seemed to take until March 13th 2020 for the UK government to decide it was not prepared to allow NHS surge capacity to be overwhelmed (and to determine precisely what that surge capacity was). In the future, it would be highly advantageous for policy planning (and science advice informing that) for such "red lines" to be clearly defined at the earliest possible stage. Furthermore, early definition of strategic policy objectives (*e.g.* minimising economic impact, minimising mortality, keeping schools open) for the response would also be beneficial, accepting that these might not all be achievable and will likely evolve over time.

461. *Healthcare capacity*: The NHS is run at substantially higher normal levels of bed occupancy and has fewer hospital beds per capita than seen in most (if not all) comparable European countries. For instance, France has twice as many hospital beds per capita as the UK. In my view, it is plausible that this difference – together with the higher peak number of hospital

admissions in the first COVID-19 wave seen in England compared with France – largely explains why France experienced a two-fold lower IFR in first wave of the pandemic than did the UK [NF/068 - INQ000262566]. Clearly, any decision to expand hospital capacity needs to consider much more than just NHS resilience to pandemic-related surges in demand, but I still think the topic requires careful consideration by the Inquiry.

N. Lessons for how epidemiological modelling informs the UK response to epidemics

Introduction

462. Epidemiological modelling is the only tool available to prospectively and quantitatively evaluate the potential impact of different policy options on the progression of an epidemic. As such, scientific advice from epidemiological modelling is often qualitatively different in its degree of immediate policy relevance (and therefore political sensitivity) than that from, say, immunology or virology (though both the latter two made substantial long-term contributions to the UK response to COVID).

463. The four key areas of evidence that epidemiological modelling can help to provide are:

a. Analysis of epidemiological and health data to characterise the threat and to evaluate (in retrospect) the effectiveness of interventions. This would include assessments of transmissibility, severity, the extent of global and national spread, the effects of viral evolution (i.e. variants) and the effectiveness of vaccines and NPIs.

b. National, regional and local situation reports and short-term ensemble-based forecasts/projections. This will include assessment of healthcare demand, testing requirements, R and growth rate estimates.

c. Longer-term scenario modelling to inform government planning and policy formulation.

d. Providing policy-makers with insights (sometimes education) into the dynamic nature of epidemics. Epidemics are complex non-linear processes, and as such lay intuition about their behaviour can sometimes be wrong. In particular, they are

not like other natural disasters which one has to "weather" for a few weeks and then get back to normal; they react to the interventions put in place.

464. Epidemiological modelling played a significant role in informing policy responses during the pandemic in many European countries. Excluding the UK, these included (but are not limited to) France, the Netherlands, Denmark, Norway, Finland, Sweden and Italy.

465. Of these, all but Italy experienced substantially lower per-capita deaths than the UK during the pandemic, whether measured by excess deaths or confirmed COVID-19 deaths.

466. My view is that the mortality the UK experienced in the first year of the pandemic would have been substantially less had rather simple conclusions from modelling and epidemiological analysis been more influential in shaping timely policy-making – particularly in the autumn of 2020 (see section J).

467. In the Nordic countries and the Netherlands, modelling was coordinated and in many instances conducted by national public health agencies (*i.e.* equivalent to UKHSA in the UK).

Epidemiological modelling in the UK

468. The UK was unique among European countries in how many academic modelling groups were involved in advising government, via SPI-M-O. Most European countries engaged a much smaller number of groups. In the Netherlands, a single institution – RIVM – conducted the modelling used to inform government decision making. In other countries, small modelling teams spanning public health agency and one or two university groups were assembled, feeding in modelling and analysis via those public health agencies.

469. In part, the size of SPI-M-O reflected the large and diverse ecosystem of epidemiological analysis and modelling in the UK. This spans UKHSA, a few large university groups (in size order: Imperial, LSHTM and Warwick) and a number of smaller groups and individual academics.

470. The distribution of academic group sizes seen in epidemiological modelling (i.e. a few large centres of excellence and many smaller groups) is similar to what is seen in most scientific disciplines. Given the highly competitive nature of grant funding panels, group size generally reflects track record as measured by scientific outputs. Whether in epidemiology or virology, size also offers some advantages in crises in terms of the scale of research which can be delivered in a timely manner.

471. Paradoxically, given some of the criticism sometimes directed at the large academic modelling groups here (Imperial most notably), modelling input into policy was less dominated by a single group in the UK than in several other European countries, even in March 2020. In my view, none of the highly consequential policy decisions made during the pandemic were solely based on the work of a single academic group.

472. Of course, as in the UK, university-based groups in European countries who were not involved in "official" modelling still published a large amount of epidemiological analysis and modelling. A minority of that research (as is always the case in science) was significant in informing scientific understanding and/or policy responses.

SPI-M

473. While SPI-M-O produced a huge amount of high quality work, often under intense time-pressure, comparing the experience of the UK with other countries suggests that this scale of activity was neither necessary nor sufficient for the UK to deliver a timely, evidence-based and effective pandemic response.

474. SPI-M has a long history – beginning in 2005 as an ad hoc "scientific advisory group" for pandemic preparedness planning in the then Department of Health. Eventually, advisory structures were formalised, and the "SPI" advisory group on pandemic influenza preparedness was created, with SPI-M as a subgroup to this. I note that I was a participant in all these groups.

475. SPI-M was never initially intended to be an operational group; to my recollection, the conception of that role ("SPI-M-O") evolved gradually between 2007 and 2009.

476. SPI-M-O was first stood up during the 2009 H1N1 influenza pandemic. A that time, it was a small group (<12 people) which included representatives from four institutions: HPA (now UKHSA), Imperial College, LSHTM and Warwick – the latter three being (then as now) the largest university-based modelling groups. At that time, both the chair (Peter Grove) and secretariat were provided by the Dept of Health.

477. SPI-M-O was stood up again to support SAGE during the West African Ebola epidemic in 2014. In that instance, the great majority of the modelling was generated by MRC GIDA and LSHTM, as both groups were supporting WHO and affected countries.

478. SPI-M outlived SPI itself, the latter being replaced by NERVTAG. However, rather than reporting into NERVTAG, SPI-M continued to report directly into DHSC. It also retained its status as an ad hoc advisory group rather than becoming a "Nolan"-governed official advisory committee like NERVTAG, JCVI or ACDP.

479. As in the COVID-19 pandemic, SPI-M-O reported into SAGE in 2009 and 2014. Unlike during the COVID-19 pandemic, SPI-M-O consensus statements then were discussed and approved (and in some cases co-written) by all participants, not just the chair(s).

The role of UKHSA

480. HPA/PHE/UKHSA has a long history of having a small but high-quality epidemiological modelling capability. Historically this was split between Colindale and Porton Down, with John Edmunds heading the former ("modelling and health economics" unit) for several years. This modelling capability informed many of the public health functions of the agency, notably vaccination programme planning. However, the capacity was never such that it was sufficient to cope with all the demands placed on it in a crisis of the scale of a pandemic.

481. During the first year of the COVID-19 pandemic, PHE provided less epidemiological analysis and modelling input into the UK policy response than I had expected (or than would have been optimal) – indeed, less than it did in 2009. Furthermore, much of the work which PHE did undertake seemed piecemeal, and lacking in central coordination. Early in the pandemic, my perception was that there was insufficient flexibility by senior PHE leadership in re-tasking the analytics capacity of the organisation away from "routine" activity towards COVID-19 work. In part this may have been because modellers and analysts were perhaps more scattered across PHE units and directorates in early 2020 than was the case in 2009.

482. Since 2020, and assisted by the formation of the JBC (and that centre's later merger into UKHSA), UKHSA analytical capability has grown considerably. Some notable highlights I have been involved with include their work on:

a. COVID-19 variants, where UKHSA coordinated external input from multiple academic groups spanning all areas of science from virology to modelling via the "Variant Technical Group".

b. Vaccine effectiveness – where UKHSA undertook world-leading studies itself, but also organised a wider vaccine effectiveness technical group spanning UK universities and research institutes.

c. MPOX – where UKHSA stood up a MPOX technical group of UKHSA and university scientists and clinicians spanning all relevant disciplines, and also invited external academic modelling groups to support and advise on UKHSA analysis and modelling of that outbreak.

d. Avian influenza (H5N1) – where UKHSA stood up a technical group to consider the human health implications of the large H5N1 epidemic in wild birds which occurred in 2022.

483. In all four of these examples, I think having epidemiological modelling considered in an integrated and detailed way at the same time as input from multiple other disciplines (including genetics, virology, immunology, field epidemiology and clinical public health) was a substantial advantage. While in theory this happens at SAGE, in reality the formality of SAGE and the time-pressures on it meant that such detailed interdisciplinary technical discussion were rarer than would have been optimal, especially given most detailed discussion of modelling occurred in SPI-M-O.

Suggestions for future governance

484. As stated earlier, of all the scientific disciplines informing the national response to an infectious disease threat, epidemiological modelling is uniquely able to give insight into the potential impacts of an epidemic on health and healthcare systems, and into how various policy interventions might modify/reduce those impacts. As such modelling will remain an important aspect of the public health response to any future epidemic.

485. However, I do not believe that SPI-M as currently conceived (i.e. reporting directly into DHSC and SAGE in its SPI-M-O form) is the best structure for the future. Its ad hoc nature, lines of reporting and disconnection from UKHSA are all problematic.

486. I recommend UKHSA be given overarching responsibility for providing "official" epidemiological analysis and modelling to guide policy-making in future public health crises.

487. PHE had always been intended to be the lead provider of modelling for UK pandemic response, and this formed part of the original conception of how SPI-M-O would operate during and after the 2009 influenza pandemic. The idea at that time was that PHE would provide the "official" view, and that other groups (*e.g.* Imperial, LSHTM and Warwick) would provide "second opinions" and/or additional analyses. However, as stated above, PHE was unable to/did not provide that function in the first few months of the COVID-19 pandemic.

488. UKHSA should have the resources and organisational structure to allow it to lead on provision of analysis and modelling to UK government during a public health crisis. It was suboptimal in the COVID-19 pandemic for universities such as Imperial College, LSHTM and Warwick to be put in that role.

489. UKHSA should also be required to work proactively and flexibly with external partners (in academia and the commercial sector) in order to fulfil this mission. It is not cost-effective for UKHSA to be permanently funded at a level that would allow them to provide *all* public health analytical needs to government in a crisis of the scale of a pandemic. The input of large and small academic groups will therefore continue to remain essential (as well as desirable). However, UKHSA should have the responsibility for organising and integrating inputs from such partners.

490. Funding work from external partners via flexible emergency mechanisms will be important; many academic groups (particularly the smaller ones) struggled to fund their COVID-19 work in the early months of the pandemic. That said, consideration should be given to what work is essential for either situational awareness or for informing policy, and what is less critical. Not everything needs to be modelled, nor are five views of every small policy question always needed.

491. A flexible range of long-term partnership models between UKHSA and academic partners should be considered, across all scientific disciplines. In addition to split appointments and secondments, more formal inter-institutional arrangements might be considered (akin to the WHO Collaborating Centre model). These should be established now, as part of UKHSA "peacetime" activities, and should be mission- rather than just pure research- oriented.

492. In fulfilling a SPI-M-O like mission in a pandemic, UKHSA should also think innovatively in terms of governance models which allow efficient scalability. I am not convinced online meetings with 70+ participants (the scale SPI-M-O reached) are an effective use of participants' time or allow sufficient detailed technical discussion and challenge. In my view, the small group (e.g. on RWC and the 2021 roadmap modelling) and sub-group (e.g. on MTPs and spatial analyses) meetings were more effective. The UKHSA Technical Group model (focussed on specific topics) has also been effective.

493. In thinking about future structures for organising epidemiological analysis and modelling, a key issue is how one ensures constructive but critical challenge. In my view this happened

to a more limited extent than would have been optimal on SPI-M-O. In part this was because of the lack of time to review and discuss individual analyses, in part because of the psychology of very large group meetings – namely in a context where everyone was working flat out, it sometimes seemed churlish or overly negative to call out issues which in other contexts (e.g. peer review of a scientific paper) might have been viewed as significant. Even when there was challenge, there was no obvious mechanism for following up; whether analyses or models were updated was generally down to individual participants.

494. I think such issues can partly be addressed by a greater focus on smaller groups – which permit more nuanced discussion, and where critical feedback is perhaps less personally challenging than in large plenary meetings. I also think it would be easier for a UKHSA-organised modelling partnership or network to be more directive or selective, in some instances.

495. A further benefit of integrating the analytics and modelling capacity provided by university groups into UKHSA capabilities is that it reduces the somewhat artificial divide between scientific advice and operational planning. Modelling (and analytics more generally) has a substantial role to play in operational planning and the evaluation of policy interventions. While UKHSA and NHSE (with additional input from technical capabilities in other departments, including the Cabinet Office) appropriately have primary responsibility for such work, efficient standing mechanisms to surge such capacity by drawing on academic (and commercial) partners would enhance future responses to major crises. This did occur during the pandemic (e.g. in modelling of test and trace, care home infection control, and of various aspects of healthcare demand modelling), but was often more ad-hoc than it might have been.

496. Epidemiological assessment of a new threat will always be a global effort in a large-scale epidemic or pandemic. Working with WHO, UKHSA should take on a more active convening role in helping to gather and synthesise research evidence from around the world. During the COVID-19 pandemic, WHO ran a global modelling network which offered a forum for the exchange of research results and resulting discussions. US CDC also ran a large network which involved some non-US members (including ourselves) and generated a very useful summary/synthesis of epidemiological parameter estimates. PHE was less engaged in these international efforts in the first months of the pandemic than was optimal, though its engagement grew over time.

497. In the COVID-19 pandemic, SPI-M-O walked a hazy line between being independent of government yet governed by it. This was illustrated by it having an independent chair, but co-chairs from within government, and a secretariat provided by DHSC (rather than by the relevant technical agency, PHE). In large part, this reflects its origins as a policy-focussed advisory committee on pandemic preparedness. Again, I think giving UKHSA primary responsibility for modelling advice relating to both preparedness and epidemic response would avoid such haziness in future.

498. This would also facilitate longer-term and more technically sophisticated partnerships between government (represented by UKHSA) and external partners, while mitigating the risk of competition (or confusion) caused by multiple sources of scientific evidence entering government at the policy/political (rather than technical) level. To a degree, PHE was disempowered from its role as the normative source of epidemiological assessment in the first few months of 2020 – even if this was partly self-inflicted.

499. None of the above precludes the need for SAGE to have external (to government) participants representing the full breadth of relevant scientific disciplines. In the case of public health crises, this will (and should) include epidemiological modelling.

500. However, I believe there are benefits in separating the roles of being an external participant in SAGE from that of being a provider/generator of policy-relevant scientific research (such as modelling). A challenge that John Edmunds and I faced on SAGE is that we undertook both roles. Indeed, in the first few SAGE-COVID meetings, I sometimes provided more information to the committee on the international situation than did PHE. Wearing two hats was not uncommon on SAGE more generally (e.g. Peter Horby was both head of the RECOVERY trial and chair of NERVTAG), but given its policy sensitivity, there are more advantages for the roles to be separated for modelling than perhaps for other disciplines.

501. As discussed in section M, it is important that SAGE is able to provide constructive challenge and feedback on modelling input it considers, as it does with other scientific evidence considered. In the COVID-19 pandemic, this challenge role was largely delegated to SPI-M-O. Indeed, all the modellers on SAGE were either co-chairs of or participants in SPI-M-O. In addition, while SPI-M-O purported to be semi-independent of government, SPI-M-O consensus statements were written by the DHSC secretariat and only signed off by the SPI-M-O co-chairs (one of whom was a civil servant).

502. I think this challenge function of SAGE would be enhanced if UKHSA were solely responsible for providing "official" epidemiological assessments in a manner which was somewhat more at arms-length from the SAGE participants considering them than was the case in the COVID-19 pandemic – recognising that complete separation of roles and responsibilities may be impractical and even undesirable.

503. A good model for the sort of epidemiological assessments I think UKHSA will need to produce in future is provided by the "Technical Briefings" they have written on COVID-19 variants, MPOX and avian influenza. These combine both UKHSA analyses and inputs from external partner academic groups into a unified assessment.

Scientific/technical lessons learned

504. I give an Imperial-focussed perspective in this sub-section. Looking back at the epidemiological analysis and modelling conducted by ICCRT throughout the pandemic, there are no major areas where I would say we got things "plain wrong".

505. Perhaps most importantly, the early estimates of disease severity (IFR) and transmissibility (R_0) we generated by mid-February 2020 have proved to be remarkably close to what the UK experienced in March 2020, especially considering how limited the data was at the time we generated those estimates. Many other academic groups generated similar estimates of R_0 at around that time, but very few (if any) accurately estimated IFR as early as ICCRT did.

506. These two estimates were significant in that they determined that in the absence of population immunity (the natural assumption to make with a new human pathogen, and one proven correct by later studies), interventions to reduce R to below 1 (a suppression strategy) were needed to avoid the UK COVID-19 epidemic from overwhelming NHS capacity many times over and leading to hundreds of thousands of deaths. This latter conclusion did not depend on the precise epidemiological model used and held true even if the IFR was less than half of what we estimated.

507. In the UK context, I would also like to call out for praise the work of our UK real-time modelling team (headed by Marc Baguelin and Anne Cori) and our genetics and genomics team (headed by Erik Volz). Both teams delivered an enormous volume of high quality work for SPI-M-O, UKHSA and SAGE, often to brutal deadlines.

508. Epidemiological modelling is rarely if ever precisely correct. All models make simplifying assumptions, and estimates of parameters going into models are subject to uncertainty. I have commented on model limitations earlier in this statement, but as more data accumulated, we revised model parameters relating to R_0, the relative contribution of children to transmission before Alpha emerged, probability of hospitalisation and ICU admissions and length of stay in both, the generation interval of COVID-19, the relative infectiousness of asymptomatics, and a number of more minor parameters. Furthermore, with the emergence of every new variant of concern (Alpha, Delta, Omicron), parameters needed to be re-estimated. While we got better at doing this, it still took several weeks in each case, largely because of the need to let data accumulate.

509. The pandemic highlighted the need for specific types of analytical tools (i.e. software) and models. While we entered the pandemic with a variety of modelling tools ranging from the simple to the complex, we invested very substantial resources in new model development during the pandemic, with a particular focus on ease of repeated use (given some were being run daily) and reproducibility. Part of that investment involved the development of automated pipelines for data processing and running models – work we had started prior to the pandemic (e.g. for the West Africa Ebola epidemic), but which evolved to a different scale during the pandemic.

510. The pandemic also highlighted the limitations of existing epidemiological models. My priorities for future research include (but are not limited to):

a. Better understanding how populations respond to epidemics, and incorporating more sophisticated representations of behaviour change into models. This is a topic I first highlighted as a priority in 2007, and while some progress has been made (e.g. in use of digital technologies to measure proxies of contact patterns), we are a long way from being able to reliably predict how people's day-to-day behaviour might change in response to, say, reports of rising mortality from a new infectious disease.

b. Better incorporation of population heterogeneity into models. The two forms of heterogeneity which I think are particularly important are (i) spatial – very few SPI-M-O models were spatially explicit in the sense that they modelled people moving from place to place (CovidSim being one exception); (ii) social – no SPI-M-O models (including our own) stratified models by clinical vulnerability (other than age), income group/deprivation

level, or by ethnicity. Yet exposure to COVID, clinical outcomes and individuals' ability to, for instance, work from home varied substantially by these factors. Technical challenges (data gaps and computational) explain this omission (see section J, last subsection), but given the importance of these factors in relation to clinical outcomes and interventions, these should be urgently addressed in ongoing research.

c. Integrating epidemiological models into models of other aspects of society – notably the economy. We have made a start on such work[51], but much more remains to be done.

Communicating epidemiological analysis and modelling

511. I and my colleagues within ICCRT actively engaged with the UK and international media throughout the pandemic (see Appendix B). This work prioritised communicating the research we were undertaking, but also included more general explanations of the developing science and of the progression of the epidemic. In this, we were supported by a dedicated communications officer in ICCRT and by the Imperial College communications team.

512. Overall, I feel that I and my colleagues were successful in communicating our research (and the science more generally) clearly and without over-simplification. I also think that modelling was well-communicated by the majority of the media, especially when done so by science journalists.

513. That is not to say I didn't make mistakes – perhaps most notably in July 2021 when I over-confidently stated it was "almost inevitable" that the UK would see 100,000 cases per day in the following few weeks (see section L). My general approach to such mistakes was to admit making them and learn lessons.

514. When communicating research findings, there can be a balancing act between focussing on explaining the potentially serious public health implications versus emphasising uncertainty and the preliminary nature of findings. This is particularly the case for scientists on advisory groups such as SAGE, where I sometimes felt a tension between communicating my analysis of the evidence and not appearing alarmist or causing "difficulties" for Patrick Vallance and Chris Whitty. To give concrete examples:

a. On February 12th 2020 I gave a BBC Today Programme interview, where I discussed our early estimates of the IFR of COVID-19 [NF/047 - INQ000228651]. For fear of being alarmist, I resisted stating outright the implication that the UK might experience 500 thousand or more deaths in the COVID-19 pandemic, though the interviewer walked

through that calculation (up to 80% of people getting infected and a 1% IFR). However, I was comfortable stating that I felt the world was then in the first stages of a global pandemic, and that I thought government border measures were detecting no more than 1 in 3 cases coming into the country. On the other hand, I don't think I explicitly stated my then belief that transmission was very likely already underway in the UK. Chris Whitty was then interviewed after me on the same programme, and while I didn't regret my early comments, I was conscious (he and I chatted afterwards) that they made his interview more difficult than it might have otherwise been. His comments in that interview were notably more cautious than mine.

b. On 10th December 2021, the Guardian published an interview with me [NF/362 - INQ000262603] where I discussed the trajectory of the Omicron wave which was taking off in the UK at that point. In that I explained the potential consequences for healthcare demand but emphasised that we needed more data on the relative severity of the Omicron variant (something I was actively trying to address at the time). While these comments were criticised by some as alarmist, in retrospect I feel comfortable that they accurately reflected the science at the time. That interview included my statement that "If it turns out, actually, it looks like hospital admissions may only peak at 2,000 to 3,000 a day, then it's possible that something like plan B – maybe a little bit plan B-plus – might be sufficient". In that context, I note that UK daily hospital admissions associated with the Omicron variant peaked at around 2900 in January 2022.

515. While this may seem rather naïve in retrospect, I also didn't anticipate quite how politically polarised some aspects of the COVID-19 pandemic would become from March 2020, and quite how much science – epidemiological modelling in particular – would get caught up with that. In January 2021, the Conservative MP Neil O'Brien correctly described much of the criticism of SAGE and SPI-M-O that had occurred in the previous summer and autumn as "motivated reasoning" [NF/363 - INQ000262590]. I also agree with his conclusion that those who peddled it – especially in the autumn of 2020 – "have a hell of a lot to answer for".

516. I directly responded to the barrage of "motivated reasoning" on occasion – for instance in a cautiously worded but in retrospect quite accurate piece for the Times RedBox site in September 2021 [NF/364 - INQ000273323].

517. In that context, I would note that over the course of the pandemic, the Telegraph, which adopted a highly "lockdown-sceptic" editorial line from March 2020 – made 15 corrections to articles criticising "Imperial" modelling after Imperial College pointed out factual errors [NF/365 - INQ000279991; NF/366 - INQ000279992; NF/367 - INQ000279993; NF/368 - INQ000279999; NF/369 - INQ000279996; NF/370 - INQ000279981; NF/371 - INQ000279990; NF/372 - INQ000279986; NF/373 - INQ000279997; NF/374 - INQ000279989; NF/375 - INQ000279983; NF/376 - INQ000279985; NF/377 - INQ000279994; NF/378 - INQ000279995; NF/379 - INQ000279984; NF/380 - INQ000279988; NF/381 - INQ000279998].

518. I believe that there were substantive consequences of this quite polarised public discourse and political environment on the management of the pandemic in the autumn of 2020, and therefore for the death toll seen in the UK.

519. More generally, long-term scenario modelling generated by a number of SPI-M-O groups played an important role in informing policy-making throughout the pandemic. To the extent that I could evaluate it (and I will be interested to learn what ministers' views were), I think the communication (e.g. by SPI-M-O and SAGE secretariats, the GCSA and CMO) of the limitations of such modelling was generally done well within government – in the sense that civil servants and ministers understood that precise long-term prediction of the epidemic was impossible, but that scenarios generated with defined parameter assumptions could be useful to inform understanding of the range of the possible and the potential impact of policy interventions. A possible exception to this positive view is how winter scenario modelling was used in the decision to move to the second England lockdown (see section J).

520. The public communication of scenario modelling posed greater challenges, in large part due to some of the politicisation discussed above. Despite being labelled as scenarios (in some cases reasonable worst case scenarios), scenario modelling was sometimes accidentally and often wilfully miscommunicated as formal prediction. A typical tactic adopted by some newspapers was to pick the highest number shown on any graph and to describe it as a prediction without context or caveats. Commonly this formed part of a wider critical narrative accusing SPI-M-O modellers (Imperial most commonly) of being relentlessly over-pessimistic doom-mongers. This became enough of an issue (in relation to modelling of Omicron) for Patrick Vallance to author an article rebutting such criticisms [NF/382 - INQ000064538] .

521. I think the formal ensemble medium-term projections generated by SPI-M-O had fewer issues around the potential for misinterpretation. The fact that they represented projections of hospitalisations and deaths made under the assumption that R would remain constant was well understood – certainly in government, but also generally in their public interpretation and communication.

522. A last important issue to consider for future preparedness is the importance of "educating" policy-makers (both civil servants and ministers) about the key science during an extended crisis like a pandemic. By educating, I don't mean explaining specific research results, but providing concise background briefings in the underlying scientific concepts which are important for policymakers to understand. Overall, I feel the Government Office for Science did a good job in this area, including the eventual provision of lectures and briefing sessions by scientists participating on SAGE and other advisory groups. These briefings spanned multiple scientific disciplines. Also, by sheer stint of being exposed to so much epidemiological analysis and modelling, I also understand that senior policy-makers gained a incrementally better understanding of infectious disease dynamics (and what modelling can and can't do) as the pandemic progressed, assisted by the GCSA and CMO and their respective teams.

O. Concluding comments

523. I do not envy the Inquiry its task, given the scope of the terms of reference and the enormous volume of material to be reviewed. However, I think the Inquiry is both necessary and valuable. I also believe it is critical for it to generate a detailed, considered and authoritative narrative of events. This will require timeliness and thoroughness to be balanced. In that context I note that this witness statement has taken me approximately two months to write, working 6-7 days a week.

524. Past public inquiries have attempted to give a definitive summary of the scientific or technical knowledge underpinning the topic or events being reviewed, through the use of panels of independent scientific experts. I am not sure this is possible (or advisable) in the case of COVID-19; all appropriate experts (whether based in the UK or overseas) were involved in some form in the response to the pandemic.

525. I also don't believe it is possible to give a definitive view on what the UK government's high level, strategic policy objectives should have been for managing the pandemic. Different elected leaders will draw different conclusions as to how to balance the very difficult ethical, public health, social and economic challenges posed by a lethal pandemic. Regardless of my own views, I never envied ministers the decisions they had to make. In addition, a broad range of public and political perspectives on how the UK should have responded to the pandemic evolved from February/March 2020. Perhaps unfortunately, I see no prospect that the Inquiry's deliberations will alter that situation – though a definitive narrative may help dampen some of the myths about events which are still prevalent.

526. However, questions could be asked about (a) how, if and when such strategic policy objectives were formulated and then adapted over time; (a) how strategic objectives (if they existed) were translated into policy/action; (c) the extent to which the policies/actions enacted were likely to achieve those objectives; and, (d) the implementation of individual policies/actions. I place emphasis on strategic objectives, as without them, policymaking tends to be reactive and overly driven by a short-term time horizon – an issue which I think affected the UK response throughout much of 2020.

527. As I have detailed above, I also think lessons can be learned about how the structures and processes the UK government uses to generate and implement policy responses to national crises. I suspect the most critical lessons relate to internal decision-making and government delivery, but I believe there are also important learnings in relation to how technical evidence is generated in to inform policymaking in time-critical contexts.

528. Last, in all my experience of working on the pandemic, I did not encounter a government official, fellow scientist or clinical colleague who was not working flat out, in challenging circumstances, to contribute as constructively as possible to the national response. This is not to say no mistakes were made. I made some, and would have done some things differently with the benefit of hindsight. I suspect that is also true for every one of the many hundreds of individuals I worked with over the course of the pandemic. But I would like to state for the record that it was a privilege to work with all of them.

P. Statement of Truth

529. I believe that the facts stated in this witness statement are true. I understand that proceedings may be brought against anyone who makes, or causes to be made, a false statement in a document verified by a statement of truth without an honest belief of its truth.

Signed: Personal Data

Dated: 11th July 2023