Witness Name:
Professor John Edmunds
Statement No: 01
Exhibits: JE/1 – JE/244
Dated: 30 August 2023
Ref: M2/SAGE/02/JE

COVID-19 INQUIRY – MODULE 2

First Witness Statement of Professor John Edmunds

I, PROFESSOR JOHN EDMUNDS, of the Department of Infectious Disease Epidemiology at the London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT will say as follows:

1. Introduction

1.1. I make this statement pursuant to the Covid-19 Inquiry's Module 2 Rule 9 Request of 9 December 2022 **('The Rule 9 Request')**.

1.2. I previously submitted a response to the Inquiry's Rule 9 Questionnaire of 2 September 2022 on 26 September 2022 **('The Rule 9 Questionnaire Response')** and a witness statement in response to the Inquiry's Module 1 Rule 9 Request on 20 January 2023 **('The Module 1 Statement')**.

1.3. The matters I set out in this statement are within my own knowledge save for where I state otherwise. Where I refer to facts not within my own knowledge, I will provide the source for those facts. The contents of this statement are true to the best of my knowledge and belief.

1.4. Where I refer to SAGE or NERVTAG meetings, and minutes of those meetings, throughout this statement, I have not provided copies of the meeting minutes, as I

expect that the relevant Secretariats will hold, and be able to provide copies of, those minutes.

1.5. I hold a Chair in Infectious Disease Modelling at the London School of Hygiene and Tropical Medicine **('LSHTM')**. I have been involved in pandemic planning at the UK level since the mid-2000s, when I was head of the Modelling and Economics Unit at the Health Protection Agency **('HPA')**, which is now known as the UK Health Security Agency **('UKHSA')**.

1.6. I was one of the first members of SPI-M **('The Scientific Pandemic Influenza Group on Modelling')**. During non-pandemic periods this committee sits within the Department of Health and Social Care **('DHSC')**, but during pandemics it is redesignated as SPI-M-O (the 'O' stands for operational) and feeds into the Government Scientific Advisory Group for Emergencies **('SAGE')**.

1.7. Although I left the HPA in June 2008 to take up my Chair at LSHTM, I continued working on pandemic influenza and continued on SPI-M, including during the 2009 H1N1 "Swine Flu" pandemic. I still serve on SPI-M. I also attended SPI-B during the Swine Flu pandemic as the 'modeller' to act as a link between the committees.

1.8. In addition, I became a member of the New and Emerging Viral Threats Advisory Group **('NERVTAG')** in 2014 (also a DHSC committee) and continued to serve on this committee until 2022.

1.9. I attended 97 SAGE meetings throughout the Covid-19 pandemic. My records suggest that I attended 99 SPI-M-O meetings and 91 other SPI-M-O related meetings (including meetings of various subgroups such as the Short-Term Forecasting and Medium-Term Projections groups). I also attended 74 NERVTAG meetings.

1.10. I participated in a number of other groups that were set up during the Covid-19 pandemic, either as subgroups of SAGE (for example, the Environmental Modelling Group and the Children's Task and Finish Working Group), the DHSC (for example, the Moonshot Scientific Advisory Group and Therapeutics Clinical Review Panel Modelling Group), Public Health England **('PHE')**/UKHSA (for example, the Testing Initiatives Evaluation Board and Variants Technical Group) and other UK Government Departments (Events Research Programme Science Committee). I was not a member of any non-UK Covid-19 advisory committees during the pandemic.

2. Models of Infectious Disease Agent Study ('MIDAS')

2.1. MIDAS is a US-based network of infectious disease modellers. I am on their mailing list and therefore receive a newsletter every few months. I have never received any funding from MIDAS, and I did not work with or for them during the Covid-19 pandemic or at any other time.

3. The Centre for Mathematical Modelling of Infectious Diseases ('CMMID') and the CMMID Working Group on Covid-19

3.1. LSHTM is subdivided into three faculties, each of which has a number of departments. The departments and faculties are the formal structures within LSHTM, covering line management, financial matters, and teaching. Each member of staff is employed within a department. For instance, my post is within the Department of Infectious Disease Epidemiology, which is in turn located within the Faculty of Epidemiology and Population Health. In addition to this, LSHTM has a number of academic centres, whose role is to facilitate collaboration between staff working on a similar topic, but who may be working in separate departments or faculties.

3.2. They do this through organising seminar series, workshops, public lectures and so on. There are currently 13 academic centres within LSHTM. A list of these academic centres can be found on the 'centres' page on the LSHTM website **{JE/01 - INQ000092643}**. They cover areas such as Antimicrobial Resistance, Climate Change, Planetary Health and Statistical Methodology. Staff and students can be a member of any number of centres. It is entirely voluntary. Most are engaged with one or two centres whose research interests most closely align with their own. However, the level of engagement with the centres can vary substantially from attending an occasional meeting to being an active member of the steering or management group. The centres do not hold grants or employ or manage individuals. These functions are taken up by the departments and faculties.

3.3. The CMMID is an LSHTM academic centre. Membership is difficult to determine as individuals may have very different levels of engagement. There are, however, 156 current members of the CMMID Slack Workspace (Slack is a messaging tool that we use for internal communications). As slack requires a small annual subscription, it gives a reasonable estimate of active CMMID members.

3.4. The Directors and management team are rotated roughly every 3 years. The Co-Directors of CMMID are currently Dr Rosalind Eggo and Professor Stefan Flasche. Professor Graham Medley was the Director during much of the Covid-19 pandemic.

3.5. The Centre is organised into a number of research themes, which can be found on the 'Research Themes' page of the LSHTM website **{JE/02 - INQ000092654}**. The research themes cover topics such as real time outbreak analytics, computation and inference and tuberculosis.

3.6. Members of CMMID collaborate with many researchers and public health agencies across the world. This includes other modelling teams. A given research project seldom requires the input of two modelling teams, who may well have similar skill sets. Hence, collaboration with modelling teams – whilst not rare – is less common than collaboration with other academic or public health teams. There are many exceptions to this, including the Health Protection Research Unit in Modelling and Economics, which is a National Institute of Health Research **('NIHR')** funded collaboration between the LSHTM, Imperial College and UKHSA modelling teams, with the aim of improving modelling methodology.

3.7. It should be emphasised, however, that despite this there was little direct collaboration with either Imperial College or UKHSA over the course of the pandemic on UK-related Covid-19-related issues. This was deliberate. The role of SPI-M is to collate and review modelling inputs from independent groups. It is important, therefore, to maintain a degree of distance from other groups feeding into this process.

3.8. It was clear by early to mid-January 2020 that the novel coronavirus outbreak in China was a major public health threat. This was likely to require significant epidemiological and modelling input to better understand and quantify the nature of this threat and to guide the response to it. As CMMID is well known as being a centre for excellence for mathematical modelling and outbreak analysis, it was increasingly obvious that members of CMMID were likely to be required to respond to this. Indeed, some were already working to this effect.

3.9. It was also, therefore, obvious that we would have to organise ourselves to better respond to the likely increase in volume, tempo, policy significance and public interest in our work. Hence, we created the CMMID nCoV working group, which was later

renamed the CMMID Covid-19 working group when this new disease was officially designated Covid-19.

3.10. The idea of the working group was to mobilise the surge capacity needed, to coordinate the work on Covid-19 to ensure that critical questions were addressed, to minimise overlap, and to improve the standard of the work through a system of internal peer-review. We also took the decision early on to recognise all members of the working group on publications arising from the work as it was not possible to accurately keep track of who was working on what and many individuals worked on aspects that underpinned several studies (e.g., those that worked on the data pipelines).

3.11. The CMMID Working Group first met on 21January 2020. We met weekly throughout the pandemic. The working group was brilliantly led and organised by Dr Rosalind Eggo, who at the time was one of the Deputy Directors of CMMID. A management team was also formed, comprising those who had line management responsibilities for individual researchers within the working group. This team also met separately every week for much of the duration of the pandemic. The working group reported to no-one.

3.12. Membership of the working group was open to anyone within CMMID. These people are mainly academics and PhD students working on infectious disease transmission. Membership was voluntary, but dependent on the agreement of line-managers. The level of engagement varied between individuals and over time. Some individuals worked full-time, others on a part-time basis. The number of actively engaged individuals fell over time, particularly as the epidemic entered its second year. However, there was no record of individuals' time spent working in the working group and so there was no way to quantify the person-hours devoted to Covid and how this changed over the course of the epidemic. Many individuals worked well over their nominal hours for many months and even years.

3.13. At its height there were over 60 individuals active within the CMMID Covid-19 Working Group. There were different sub-groups within the working group, which tended to work on different topics, but these groups did not have fixed memberships and there was considerable overlap between them. COVID-M was developed by Dr Nick Davies, initially, with assistance from Dr Rosalind Eggo, Professor Mark Jit, myself and others. Later Dr Rosanna Barnard was brought on board and took the lead for much of the work during 2021 and beyond.

3.14. There were different teams working on different aspects of the pandemic in low to middle income countries, including Dr Carl Pearson, Mr Kevin Van Zandvoort, Dr Chris Jarvis, Dr Nick Davies, Dr Roz Eggo, Professor Stefan Flasche, Professor Mark Jit and Professor Francesco Checchi, who modelled a range of potential scenarios and the impact of alternative interventions for many different countries. Dr Matthew Quaife and others looked at the impact of non-pharmaceutical interventions in Kenya. Dr Carl Pearson, Dr Tim Russell and others quantified the transmissibility of the 501Y.V2 variant (later known as the Beta variant) in South Africa, and later worked on quantifying the infectiousness and degree of immune escape of the Omicron variant by careful analysis of South African data.

3.15. Others, including Dr Simon Proctor, Dr Fiammetta Bozzani and Professor Anna Vassell assessed the cost-effectiveness of Covid-19 vaccination in low-income settings. Another group, led by Professor Francesco Checchi, looked for signals of Covid-19 transmission by assessing burials in low-income settings such as Mogadishu and Somalia. The names listed here are just a sample of those individuals who were working on these projects, and there was often significant overlap in the groups, with individuals working on multiple projects simultaneously and others who played small but important roles preparing data, inputting on methods and reading and editing the manuscripts.

3.16. Scientists normally release their findings in academic journals following review by independent referees who work within the field. This system of peer review is designed to improve the standards of scientific articles and weed out those that may be poor quality. This process takes time – typically months to years. The urgent requirement for information on this new disease meant that the peer review system was often circumvented by the scientific community during the Covid-19 pandemic, with "pre-prints" of papers being released onto organisational websites, or servers such as MedRXiv before being peer-reviewed.

3.17. This practice had the potential to undermine the normal quality control mechanisms for scientific output. To ensure the quality of CMMID's output during the pandemic, and enable rapid dissemination of results, the CMMID Working Group instigated a system of internal peer review for all work that was released into the public domain. Individuals (usually around 6) within the working group who had not worked on the paper in question were asked to review and comment on the finished draft within a set time limit

of two working days. Later in the epidemic, the working group also instigated a system of code-review, whereby computer code related to an output was also reviewed in a similar way. Again, this system was put in place to try to ensure the high quality of CMMID output.

3.18. In addition, to ease discovery of CMMID outputs, papers were uploaded to the CMMID repository. The CMMID repository can be found online, and a copy of the webpage is exhibited as **{JE/03 - INQ000092665}**. They were often then posted on a pre-print server (often MedRXiv) and then submitted for peer-review publication through the traditional journal route. The CMMID repository was launched one week after the first Working Group meeting (on 28 January 2020). It contains hundreds of reports and papers produced by the working group to help inform the scientific response to the pandemic.

3.19. Interactive applications (knows as 'apps') were also added to the repository when appropriate and resources allowed. These were associated with papers and enabled users to interact with the models directly – changing the parameter values to those that suited their situation or interests best and exploring the impact of these changes on the results. A number of such apps were uploaded to the CMMID repository. At present, the apps featured on the CMMID repository website are not functioning. The apps covered issues such as hospital demand forecasting, simple ways to infer the size of the epidemic from data on deaths, the effectiveness of airport screening, an online version of our main transmission model (COVID-M), and summaries of CMMID work done in different areas of the world (e.g. the Middle East and North Africa) and Brazil.

3.20. A huge range of work was undertaken by the CMMID Working Group over the course of the pandemic. It is impossible to adequately summarise it here. Indeed, this is exactly why we set up the online CMMID repository, so that interested individuals could easily browse or search for the work that was undertaken. I offer up a few highlights below to illustrate the range of analyses that were performed by the CMMID. All of this work is published on the CMMID repository website, and I have included links to some of the studies conducted, and links to summaries of some of the studies, as exhibits to this statement. Should the Inquiry wish to view full versions of the summarised studies they can be accessed on the CMMID repository website.

4. The work of the CMMID and the CMMID Working Group

4.1. At the outset of the epidemic, we concentrated on understanding the nature of the initial outbreak in Wuhan and characterising the new disease in terms of key epidemiological quantities, such as the basic reproduction number (R0) and serial interval distribution (the time between generations of cases). We looked at simple interventions that were likely to be considered early in the epidemic, such as travel restrictions and contact tracing, and initiated a host of other work that would begin to bear fruit later, including starting to develop and parameterise spatial and age-structured models of transmission in the UK and elsewhere.

4.2. By the end of January 2020, we had uploaded on to the CMMID Repository (and submitted for publication in most cases) papers that assessed the nature of the initial outbreak in Wuhan and estimated that R_0 was likely between 2 and 3 **{JE/04 - INQ000092676; JE/04A - INQ000255393}**.

4.3. We estimated the reporting delays in China, which is critical for understanding the data. **{JE/05 - INQ000092687}**. In estimating the reporting delays in China, we came up with a realistic real-time assessment of the number of cases in China (which was far bigger than the reported cases).

4.4. We analysed multiple data sets simultaneously, including those from individuals who had left Wuhan (or been repatriated) to provide a rigorous assessment of the scale of the epidemic there, the potential for global spread and the impact of interventions (lockdown) on transmission in China **{JE/06 - INQ000092695; JE/06A – INQ000255394}**. We assessed the value of airport screening, and a summary of this work is also found on the CMMID Repository **{JE/07 - INQ000092696; JE07A – INQ000255395; JE07B – INQ000255396; JE07C – INQ000255397; JE07D – INQ000255398}**.

4.5. By the first week of February, we had uploaded papers on the extent to which transmission may occur before symptoms are apparent **{JE/08 - INQ000092697}**. We produced a simple method for estimating how many cases there are from the number of deaths **{JE/09 - INQ000092698; JE/09A – INQ000255399; JE/09B – INQ000255400; JE/09C – INQ000255401; JE/09D – INQ000255402; JE/09E – INQ000255403; JE/09F – INQ000255404; JE/09G – INQ000255405}**. This was a method which was very useful for situational awareness at the beginning of the epidemic, when very low and variable rates of testing meant that the case numbers were extremely unreliable and gave a very optimistic view of epidemic progress. This simple method was quickly refined and updated **{JE/10 - INQ000092644}**.

4.6. We conducted an early, yet rigorous, assessment of the feasibility of controlling Covid-19 through contact tracing and isolation **{JE/11 - INQ000092645; JE11A – INQ000255406}**. This was quickly followed by an assessment of the possible impact of targeting air travellers at slowing the epidemic **{JE/12 - INQ000092646; JE/12A – INQ000255407; JE/12B – INQ000255408}**. A number of these papers were highly influential, as measured by the number of times they have been cited in the scientific literature, including the study on contact tracing (at **{JE/11 - INQ000092645; JE/11A – INQ000255406}**), which I have been asked to comment on.

4.7. During February and early March 2020, we continued to build our knowledge of Covid-19 through analysis largely of the international data, including refining our estimates of the case-fatality ratio (what proportion of cases die) through analysis of the outbreak on the Diamond Princess cruise ship that had been quarantined off Japan **{JE/13 - INQ000092647}**.

4.8. By early to mid-March 2020, we were able to put all these aspects together to come up with assessments of the impact of different control measures in the UK. A series of reports were sent to SPI-M and SAGE and are summarised in the paper 'The effect of non-pharmaceutical interventions on Covid-19 Cases, deaths and demand for Hospital Services in the UK: A modelling Study' **{JE/14 - INQ000092648; JE/14A – INQ000255409; JE14B – INQ000255410}**. We also developed the tools to track and forecast the epidemic progress in countries with reliable data **{JE/15 - INQ000092649**; **JE15A – INQ000255411**; **JE/16 - INQ000092650}**.

4.9. By the end of March 2020, we were simulating what the impact of the epidemic might be in LMICs **{JE/17 - INQ000092651; JE/17A – INQ000255412; JE/17B – INQ000255413; JE/17C – INQ000255414}** and had come up with the first quantitative estimate of the impact of the lockdown in the UK on the reproduction number **{JE/18 - INQ000092652**; **JE/18A – INQ000255415}**.

4.10. The latter part of the first wave saw us refining estimates of key parameters **{JE/19 - INQ000092653; JE/19A – INQ000255416}** integrating these estimates into updated analyses on a range of different measures in many different countries, **{JE/20 - INQ000092655}** including estimating the global distribution of high-risk individuals **{JE/21 - INQ000092656; JE/21A – INQ000255417}** and investigating in further detail the likely performance of track, trace and isolate programmes **{JE/22 - INQ000092657**; **JE/22A – INQ000255418**; **JE/23 - INQ000092658; JE/23A – INQ000255419}** as well as travel restrictions **{JE/24 - INQ000092659**; **JE/24A – INQ000255420**; **JE/25 - INQ000092660; JE/25A –INQ000255421}**.

4.11. The upswing of cases during the summer that accelerated into the autumn of 2020 saw us assess the impact of different control policies, including further lockdowns of different duration and intensity **{JE/26 - INQ000092661; JE/26A – INQ000255422; JE/26B - INQ000255423; JE/26C - INQ000273554}** as well as assess the impact of the tiered restrictions that were in place in the UK at the time **{JE/27 - INQ000092662; JE/27A – INQ000255424}**. There was also the prospect that testing would become more generally available, particularly through the provision of lateral flow tests, and a series of papers were produced looking at the likely impact of expanding testing to improve isolation and quarantine policy **{JE/28 - INQ000092663**; **JE/28A - INQ000255425; JE/29 - INQ000092664**; **JE/29A - INQ000255426**; **JE/30 - INQ000092666; JE/30A - INQ000255427; JE/30B - INQ000255428}**.

4.12. We looked at the impact of mass testing on the epidemiology as a whole **{JE/31 - INQ000092667**; **JE/31A - INQ000255429}.** There was also the prospect that vaccination would be on the horizon, and we assessed the cost-effectiveness of these policies **{JE/32 - INQ000092668**; **JE/32A - INQ000255430}**.

4.13. In December 2020, the first of the new variants were discovered, here in the UK (later designated the Alpha variant) and in South Africa (later designated Beta). The rapid increase in cases required an equally rapid re-assessment of the transmission and immune escape characteristics of these viruses **{JE/33 - INQ000092669**; **JE/33A – INQ000255431**; **JE/33B - INQ000273555; JE/33C – INQ000255432; JE/33D – INQ000255433; JE/33E – INQ000255434**; **JE/34 - INQ000092670; JE/34A – INQ000255435**; **JE/34B – INQ000255436}** as well as an assessment of the impact of tightened interventions to control their spread. In January 2021 we noticed that the estimated number of deaths from COVID-19 was higher than we had predicted, which stimulated us to investigate further by undertaking a detailed statistical analysis of a large-scale linked dataset within the UK. This led us to discover that the Alpha variant was also significantly more pathogenic than the previously circulating strains of COVID-19 **{JE/35 - INQ000092671; JE/35A – INQ000255437}.**

4.14. The spring and summer of 2021 saw a series of studies looking at the gradual easing of restrictions (the roadmap out of restrictions). These are publicly available and mentioned at **{JE/36 - INQ000092672}**. There was also work on the characterisation of the Delta virus **{JE/37 - INQ000092673; JE/37A – INQ000255438; JE/37B – INQ000255439; JE/37C – INQ000255440; JE/37D – INQ000255441}**, the role of re-infection **{JE/38 - INQ000092674; JE/38A – INQ000255442}** and ongoing assessments of the possible impact of vaccination **{JE/39 - INQ000092675; JE/39A – INQ000255443}**.

4.15. November and December of 2021 were dominated by a rapid assessment of the characteristics of the Omicron variant and its potential impact in the UK **{JE/40 - INQ000092677; JE/40A – INQ000255444; JE/40B – INQ000255445; JE/40C – INQ000255446}**.

4.16. The role of children in transmission and the importance of school closure, testing, or (later) vaccination of children as a means of control was difficult to ascertain due to common occult infections in younger age groups. These were topics that we returned to frequently during the course of the epidemic **{JE/41 - INQ000092678**; **JE/41A – INQ000255447**; **JE/42 - INQ000092679**; **JE/42A – INQ000255448**; **JE/42B – INQ000255449; JE/43 - INQ000092680**; **JE/43A – INQ000255450}**.

4.17. Throughout the epidemic we measured contact patterns as a means of obtaining rapid estimates of the impact of different non-pharmaceutical interventions and assessing individuals' attitudes to risk – details of this will be covered in the section on the CoMix study. The weekly and other reports can be found on the page titled 'Mixing Patterns' on the CMMID Repository webpage **{JE/44 - INQ000092681}**.

4.18. It is important to note that the vast majority of this work was not done in response to a direct request from a government department or public health agency. It was self-directed, often done in anticipation of likely future policies, or for scientific interest (quantifying the characteristics of a novel disease).

5. CoMix Social Contact Survey

5.1. Transmission of close-contact infectious diseases, such as measles, influenza and SARS-CoV-2 (the virus that causes COVID-19) are spread when an infected individual comes into contact with a susceptible individual. Thus, human contact patterns – how many contacts are made, with whom and what the nature of contact is, are major determinants of the spread of these infections. Understanding and quantifying these contacts can therefore greatly improve our understanding of the epidemiology and can greatly improve the validity and predictive power of mathematical models.

5.2. The first large-scale attempt to try to quantify epidemiologically relevant contact patterns was the POLYMOD study which was conducted in eight European countries, including Great Britain. This study was led by myself, and the summary paper from that project is publicly available on the PLoS Medicine website **{JE/45 - INQ000092682}.** It remains the most highly cited paper of my career with around 2,800 citations in the scientific literature since it was published in 2008. This can be seen on my profile on the google scholar webpage **{JE/46 - INQ000092683}**.

5.3. The study not only quantified how many contacts that an average person made, but also with whom (e.g. children tend to contact children of a similar age, and adults also mostly contact adults), and where these contacts occurred (at work, home, whilst socialising etc). These data have therefore informed a multitude of infectious disease models over the years since they were collected and have spawned a number of similar studies in different countries **{JE/47 - INQ000092684}**, including two more in the UK. These are The Warwick Contact Survey and the BBC Pandemic Contact Survey **{JE/48 - INQ000092685**; **JE/49 - INQ000092686}**.

5.4. The data from these studies (particularly the POLYMOD study) form the bedrock of our mathematical models of close-contact transmission, as they give a quantitative description of epidemiologically relevant behaviour.

5.5. However, it was known that during a pandemic of a highly infectious and pathogenic disease, individuals may change their contact behaviour to try to reduce their risk. Indeed, it was possible that governments would try to alter contact behaviour – either through the provision of public health advice, or through mandated orders. Models relying on pre-pandemic contact patterns would therefore be inaccurate if individuals did, indeed, change their behaviour during the pandemic. For this reason, in February 2020 as a matter of urgency we applied for funding for a pandemic-specific contact survey.

5.6. Our plan was for this survey to be repeated regularly to provide an up-to-date assessment of contact patterns. We applied for two grants, one to the European Union Horizon 2020 scheme (as part of a wider grant on modelling COVID-19 called EPIPOSE led by the University of Hasselt) the other grant was to UK Research and Innovation ('UKRI') as part of wider support for LSHTMs modelling efforts. The EU funding supported 8 surveys of around 1500 individuals in size every 2 weeks in three countries (the UK, Belgium and Netherlands) and the UKRI provided similar funding for the UK.

5.7. As both grants were funded, this allowed the UK survey to occur weekly – with two panels recruited, each reporting every two weeks. The survey was launched in the UK on 24 March 2020 (the day after lockdown was announced). The Belgian and Dutch surveys were launched in April 2020. The data proved to be invaluable for understanding and tracking the impact of non-pharmaceutical interventions and so the UK survey was refreshed and expanded (to roughly 3000 participants per week) in August 2020 via a follow-up grant from the UKRI and further follow-up funding was awarded from NIHR and then UKHSA so that the survey continued collecting weekly data until early March 2022.

5.8. We also made the questionnaire available to a number of other countries and Norway, Germany and New Zealand each launched their own versions in Spring 2020. Although the survey was representative of the UK population, the sample sizes in the devolved nations of the UK were relatively small. Hence, Northern Ireland and Scotland also launched their own CoMix surveys in 2020 and continued to collect boosted samples from their own populations over the course of the pandemic. The European Centre for Disease Control ('ECDC') also recognised the importance of these data and instigated and facilitated the expansion of the European aspect of CoMix, which resulted in the EU Horizon 2020 grant being expanded so that data could also be collected from another 17 countries across the EU and Switzerland.

5.9. In each country, a representative sample of adults over the age of 18 years (initially) were recruited into the survey, via a quota sampling technique (with quotas based on age group, gender and region). The survey company Ipsos was appointed to run the field work in the UK and in the other countries that were funded by the EPIPOSE grant.

5.10. In May of 2020 the UK sample was expanded so that parents of children under 18 years of age were also recruited and asked to complete the survey on behalf of their children (roughly 500 parents per week). Details of the UK survey and results over the first year of CoMix in the UK are given in Gimma et al. which is available on the PLoS Medicine website **{JE/50 - INQ000092688}**. Details of the weekly reports and other papers that were produced are on the page titled 'CoMix Survey-Social Contact Study' on the CMMID Repository **{JE/51 - INQ000092689}**.

5.11. As we had baseline data on pre-pandemic contact patterns (from the POLYMOD and other studies) we could quantify the impact that the initial lockdown had on contacts. In addition, these data allowed us to quantify the impact of the lockdown on the reproduction number 3-4 weeks before methods using the epidemiological data could do so accurately **{JE/52 - INQ000092690}**. The survey, therefore, gave the first quantitative estimate of the effectiveness of the lockdown and the first indication that the reproduction number was likely to be significantly below one because of the restrictions put in place in March 2020. These early results were immediately reported to SAGE (on 31 March 2020).

5.12. Other aspects of risk-reduction were also measured by CoMix, such as the adoption of mask-wearing, as well as some indication of the drivers of changes in behaviour (e.g. whether individuals perceived themselves to be at high risk of severe disease) **{JE/53 - INQ000092691}**. The survey therefore provided an invaluable insight into the effectiveness of NPIs in close to real-time as well as providing input to models.

5.13. There were limitations to the survey. First, children were not asked about their contact patterns directly – instead parents were asked to do this on behalf of a child in the household. Second, the survey was longitudinal in nature (a deliberate design feature) with individuals recruited for a maximum of 10 survey rounds. However, there was evidence of survey fatigue, such that individuals reported fewer contacts with increasing numbers of survey rounds. To reduce the effect of this, we increased the refresh rate of the survey. In addition, although the survey was based on the POLYMOD study, the questions were not identical and the CoMix survey was considerably longer, which may have reduced the comparability of the results with the baseline data.

5.14. For most of the epidemic, we produced a weekly report of CoMix. These were circulated every week to SPI-M, SAGE, CSA, the Cabinet Office, DHSC, and PHE/UKHSA. The data were also made available to NHS England for use in their model. These weekly updates typically reported the mean number of contacts by age group and setting (home, work, other) or geographical area (region or country of the UK) and how this was changing over time, as well as other key indicators such as the fraction of individuals currently in isolation or quarantine (by age group), usage of face-coverings and patterns of work / school attendance (and associated contact rates).

5.15. The data were usually discussed briefly at the weekly SPI-M meetings and frequently at SAGE (particularly after large policy changes). In addition to the weekly reports, periodic additional reports were produced, looking at: the impact of the tiered restrictions, the possible impact of schools re-opening, and the change in contacts that occurs over the Christmas period **{JE/54 - INQ000092692}**. These weekly and specialised reports were intended to help guide decision-making.

5.16. Other specialist reports were aimed at a technical audience, such as quantification of age-contact matrices for use in transmission models **{JE/55 - INQ000092693}** and an assessment of the likely variation in secondary case distribution (i.e., the likelihood of super-spreaders) from variation in contact and viral load data **{JE/56 - INQ000092694}**.

5.17. In addition to the UK-focussed reports (of which there were about 100 produced during the pandemic) we also produced reports for the other 17 European countries participating in CoMix. In addition, a series of other papers comparing contact rates across Europe were also produced, many of which are still going through the peer-review process at the time of writing this statement.

6. SAGE and its sub-groups

6.1. I had previously attended SAGE during the West African Ebola crisis in 2014-16. I gained considerable additional experience of SAGE and its subgroups during the COVID-19 pandemic. During the period of interest for the Inquiry I attended 97 meetings of SAGE, 99 SPI-M-O meetings and 74 NERVTAG meetings. Two of these subgroups were in existence before the pandemic, (i.e. SPI-M (the pandemic influenza modelling subgroup and the NERVTAG), which were tasked with influenza pandemic planning and respiratory viral threats risk assessments, respectively. During the COVID-19 pandemic, these committees expanded their membership and changed their reporting lines (feeding into SAGE, rather than the Department of Health and Social Care). SPI-M also changed its focus from planning for a pandemic of influenza to providing real-time epidemiological assessments and projections of COVID-19. In doing so, its name changed to SPI-M-O (for operational). These groups often spawned further working groups, either temporary groups tasked with preparing a report on a certain topic or more permanent groups which met regularly. For instance, SPI-M-O created subgroups to work on short-term forecasting and medium-term projections, which I attended, and which met weekly during the pandemic. According to my diary, I attended a total of 91 other SPI-M-O sub-group meetings over the course of the pandemic. SAGE also acted in a similar way, with occasional short- term working groups commissioned to produce a paper on a particular topic or longer-term subgroups that met regularly. I was a member of two of these longer-term SAGE

subgroups, namely the Environmental Modelling Group Transmission Subgroup and the Task and Finish Group on the Role of Children in Transmission. Other government departments and agencies also set up scientific advisory groups on particular topics that did not necessarily have a formal reporting route into SAGE. For instance, I was also a member of the Events Research Programme Science Committee (the Events Research Programme was set up by the Department of Digital, Culture, Media and Sport); The Project Moonshot Scientific Advisory Group (Department of Health and Social Care); and two PHE/UKHSA-led committees, namely the Variants Technical Group **('VTG')** and the Testing Initiatives Evaluation Board **('TIEB')**.

6.2. There were many other committees and subgroups involved in the pandemic response. I have only listed the ones that I was directly involved with. The co-ordination of this work was a very large undertaking that was done by the Civil Service. Secretariat support was provided by the respective organisations who ran the various committees. For example, GO-Science for SAGE, DHSC for SPI-M and PHE/UKHSA for NERVTAG. I did not take my own notes for any of the meetings that I attended of the groups that I participated in during the pandemic, and in preparing this statement I have relied on the official notes that were made.

6.3. I attended all of the above groups, including SAGE, in my capacity as an expert in epidemiology and, in particular, in the design of control programmes against infectious disease. I did not represent any organisations or LSHTM itself. Other members of the CMMID also attended many meetings of the various groups over the course of the pandemic. They mostly attended SPI-M-O meetings. They also attended in their capacity as individual experts in their respective fields. As individual experts we were not designated a specific role within SAGE or in any of the other groups, this meant that we were free to contribute to discussions that may have been outside our specific areas of expertise. However, I think that this represented an important function in the process of providing scientific advice to the government. The opportunity to respond to challenge from well-informed colleagues who are not necessarily expert in a particular field can be very valuable in the refinement of the scientific advice. The Chairs of the various SAGE subgroups, such as SPI-M-O or SPI-B (or sometimes the nominated Deputies of a Chair), did have more defined roles on SAGE, as they reported back to SAGE on the work of their respective subgroups. However, they were also free to contribute to other areas outside their designated roles, as ordinary

attendees did. I never thought that this was a problem – possibly because of the quality of the Chairing of the committee.

6.4. SAGE received information from its subgroups on a regular basis. The Chairs of the various subgroups attended SAGE, and some had a regular agenda slot to update the committee. For example, there was a modelling consensus statement delivered by the SPI-M-O Chair at every SAGE meeting. Other committees that were not formally subgroups of SAGE (for example, the Variants Technical Group) also provided updates to SAGE when required. These were usually delivered by the respective Chair of the committee. Furthermore, given the wide membership of SAGE, it also received information from many other sources. This included the work of other, non-SAGE aligned committees, via ordinary members of the committees in question. Overall, SAGE was very well-informed. It ensured that it had access to the latest scientific data and incorporated this into its advice. I cannot think of an example of when a significant scientific advance anywhere in the world was missed by SAGE. That is, information was rapidly assessed and assimilated into its scientific advice. This was an enormous and successful undertaking. It was aided by the fact that a number of the key committees, such as SPI-M and NERVTAG, were already in existence at the outset of the pandemic and had established networks and ways of working as well as a plan for how they would operate in an emergency. From speaking to international colleagues, my impression is that the SAGE system was well-regarded, not just in terms of the quality and efficiency of the scientific evidence that it produced, but in its openness and the speed with which it was established.

6.5. As there may be multiple interpretations of a given piece of evidence, it is possible that the consensus approach adopted by SAGE slowed the publication of clear advice. Where there was uncertainty – perhaps arising from different interpretations of some data – then this would be reflected in the uncertainty statement attached to the advice (standard Government-endorsed phrases were used). In mitigation, SAGE and its subgroups met very regularly. In the early part of the epidemic SAGE met twice weekly (as did SPI-M). For most of the rest of the period, it met every week. This helped ensure that emerging evidence was relatively quickly assimilated into advice.

6.6. SAGE was co-chaired by the Chief Scientific Advisor **('CSA')**, Sir Patrick Vallance and Professor Chris Whitty, the Chief Medical Officer **('CMO')**. They were responsible for chairing the meetings and reporting SAGE's findings to central government.

Representatives from government departments did attend SAGE. Indeed, the Departmental Chief Scientific Advisors were active participants. Other government departmental representatives, such as senior civil servants, sat in as observers. SAGE participants did not have direct contact with senior politicians. However, the Prime Minister did send an observer, Dr Ben Warner, to SAGE meetings from early February 2020 onwards and Dominic Cummings also attended some SAGE meetings during March and April 2020. The CSA and CMO were scrupulous about not reporting back discussions held within central Government to SAGE. That is, the information flow was one-way and the distinction between SAGE, which discussed and summarised scientific evidence, and policy-making or decision-making bodies, was maintained throughout the pandemic. It is impossible to ascertain whether a different model such as joint meetings between decision-makers and SAGE would have resulted in qualitatively different, or more rapid, decisions as we did not try out another system. Hence, any inferences about its effectiveness are tentative at best. Nevertheless, I have a few reflections of my own on this aspect of the process. They are:

6.6.1. It was important to keep a distinction between scientific advice and policy-making. Decision-makers had to weigh-up scientific advice alongside other evidence, for instance on the possible economic and social impact of different policy options. These were extremely difficult decisions, with enormous consequences for individuals across the country, indeed, there were life and death consequences. As a scientific advisor I am not mandated to play any direct role in these decisions, nor would I have wanted to. It is the role of Government to take these decisions – that is the basis of our democratic system.

6.6.2. The one-way flow of information was important to maintain, as it allowed the Government to discuss these difficult matters in private. However, it did have consequences. At times it was difficult for us to ascertain what the limits of Government action would be, perhaps resulting in conservativism with regards the interventions modelled, particularly in the early phase. For example, we were slow to model the implication of lockdown policies in detail (as explained later). It was not clear to me that such radical measures were politically acceptable and so we did not spend as much time on them in early March as we should have. Later in the pandemic (from December 2020 onwards) the COVID-19 Taskforce was re-organised and was very ably led by Rob Harrison. He used to attend SPI-M-O and

SAGE meetings, and although he was also strict about maintaining confidentiality, his active presence in these meetings allowed a better understanding of what could be conceivable policy.

6.6.3. It is possible that in the process of summarising the scientific evidence for the government, that some of the nuance or sense of urgency could have been lost. The minutes of scientific meetings tend to be rather dry and the early SAGE minutes were very sparse. Some of the scientific discussions were highly technical. It is not clear how much decision-makers would have benefited from being exposed to such discussions, particularly when their time was also very precious.

6.6.4. It is also almost certainly better for public confidence in the process that scientific advice is given – and seen to be given – independent of political interference.

6.7. Overall, I think that the separation of scientific advisory committees and government decision-making was important. Indeed, there were times, such as the joint Downing Street press conferences, or the often repeated "following the science" mantra, when the distinction between advisors and decision-makers might have appeared to be blurred. This ran the risk of damaging the integrity of the scientific advisory process.

6.8. The role of the committees, (for example SAGE and NERVTAG) was to discuss and review scientific evidence. The committees often took on the task of summarising their evidence in a review paper. Small subgroups were usually identified to help write such a paper, which would be brought to the main committee and further discussed there. SPI-M-O also followed this process, but its way of working was distinct from the other committees in a number of important ways.

6.9. First, it also undertook a number of routine tasks that it completed every week, including estimating the reproduction number and conducting short-term projections. These required different groups around the country to update their analyses every week and to send their analyses to a central group based at the Defence Science and Technology Laboratory **('DSTL')**, who then formally combined these analyses into a statistically rigorous consensus view. These ensemble estimates and projections, and the different analyses that had contributed to them, were discussed at the relevant subcommittee meeting of SPI-M-O every week, and then again at the main SPI-M-O meeting. These were then signed off and taken to SAGE.

6.10. In addition to these routine analyses, SPI-M-O also commissioned bespoke analyses on a regular basis. They did this on most weeks. Again, individual research groups would bring their analyses of the specific commission to the next meeting, where they were discussed and a consensus statement on that topic was then written. The aim was to harness the power of the different modelling groups around the country and to bring a diversity of methods and analyses to any given topic as a means of external validation or cross validation. Differences in the results of these separate analyses served to inform the level of confidence in the consensus statement.

6.11. In addition, much of the work that SPI-M-O discussed was not commissioned. Individuals or groups also undertook self-directed research and brought their analyses to the committee for review and scrutiny. These analyses sometimes resulted in a request to other modelling groups to see if they also found similar results. A good example of this was the work that Dr Nick Davies (of CMMID, LSHTM) undertook on increased risk of death resulting from the Alpha (as it became known) variant **{JE/57 - INQ000212186}**. This work was discussed in detail at SPI-M-O in January 2021. Other modelling groups contributing to SPI-M-O then quickly undertook their own analyses, confirming these initial results.

6.12. Thus, SPI-M-O was distinct in that its participants were active every week in conducting novel analyses (i.e. new primary research). This was assessed and assimilated into the weekly consensus statements. Given the uncertainty inherent in any real-time analysis or projections (where these were done) the committee played a key role in reviewing the evidence, corroborating it, and summarising it in an appropriate consensus statement (or other product, such as R estimates) that reflected the level of uncertainty in the analyses.

6.13. At all times SPI-M-O tried to ensure a plurality of independent analyses were conducted. During the early part of the epidemic this was more difficult to achieve, with the two largest modelling groups (at Imperial College and at LSHTM) being quicker to develop the tools and teams necessary to take on the enormous number of tasks that were required. At this critical time in the epidemic, being January to March 2020, there was therefore less cross-validation of results than would have been ideal.

6.14. At the outset of the epidemic, membership of SAGE and the various sub-committees was relatively small. This changed during the Spring of 2020, with SAGE membership,

the number of sub-committees and the membership of these committees expanding enormously. The switch to online participation, that occurred at lockdown, almost certainly helped expand the membership as the physical difficulties of fitting people into conference rooms disappeared. This also helped expand the geographical reach of the committees as the difficulties of attending a meeting in London would be negated. It is worth pointing out that before the lockdown in March 2020 SAGE and its subgroups were meeting in person. Online access via a phone-in was available, but the quality of this link was very poor.

6.15. There has been some criticism that SAGE did not include sufficient clinical or public health expertise. I assume that this criticism is aimed at the early SAGE meetings. There are a number of aspects to this. First, SAGE was a scientific committee. It was not tasked with implementation or operational activities, and the Chairs were careful not to stray into these areas, as they would be the domain of other organisations, such as the NHS. Second, Public Health England (later UKHSA) was always represented, thus there was public health expertise present. Nevertheless, the line between scientific and operational issues is not always distinct and I did think, during the early phase of the epidemic, that greater input from front line organisations could have helped in terms of context and understanding of their plans. I remember being relieved when Professor Steve Powis (National Medical Director of NHS England) first attended SAGE towards the end of February 2020. As SAGE expanded, we started to receive regular clinical input from Professor Callum Semple and the CoCIN study, which I felt was very valuable, as well as a huge range of input from different experts across the spectrum.

6.16. Challenge within committees such as SAGE and its sub-committees is essential but must be handled carefully. I have tried to highlight how SPI-M-O operated by deliberately building in challenge via tasking different groups to undertake independent analyses on the same topic. Elsewhere, such challenge was not so embedded. It is certainly possible that the consensus-building approach by which SAGE and its sub-committees operated, which is very helpful for decision-makers, might have led to reduced challenge and an element of "group-think". Perhaps having a second team (or committee) to consider major issues could have helped, though I suspect that this would have put an intolerable load on those providing the primary evidence (e.g. SPI-M-O) who were already working to exceptionally tight deadlines, and led to confusion

and delay. The fact that SAGE's evidence was publicly available did lead to challenge from outside from many sources, including academics, the media, politicians, and the general public. It is more of an issue of whether there was enough challenge from within the process. The Chairs of the committees were critical in ensuring that dissident voices were heard. It is very hard to know how well this was achieved. I did not ever perceive a problem, but that is not to say that everyone felt the same way.

6.17. Many people would be surprised to learn how narrow SAGE's remit was. It was constituted to provide scientific evidence to support government decision-making. It did not consider operational matters, nor did it consider economic and other wider impacts. The phrase "following the science" was used repeatedly by government ministers, particularly during the early stages of the pandemic. Following the science would imply ignoring these other operational and economic factors. It is also oversimplistic to assume that there was one "science" to follow, as it implies that there was no uncertainty inherent in the scientific view. It is inconceivable to me that the UK Government did ignore operational and economic factors and focus solely on the science. Therefore, this phrase was misleading. The Government should never have "followed the science" alone. Nevertheless, it is highly likely that the weight of scientific evidence for a course of action was often far greater, in terms of quality and quantity, than the weight of economic evidence. This is not a failure of SAGE and its subcommittees, and the answer was not to collect less scientific evidence, but it was a failure to also harness the power of economic analyses available within the country. SAGE evidence was relevant, timely, high quality and available to all. Projections from SPI-M-O were open to scrutiny and were picked over by academics, the press and public alike. This, along with the inbuilt challenge that SPI-M-O maintained, helped ensure high standards of output. If there were projections of the economic impact of policies, then they were not open to any public scrutiny. I am unsure whether these projections were even done, and if they were, why they were not open to public scrutiny and whether they would stand up to public and professional review. If it was the case that the Government relied too heavily on the work of SAGE and its subgroups, then it is a question for other government departments to answer for not providing similarly high-quality research on the wider impacts of the pandemic and pandemic-related policies.

6.18. In working on SAGE and its sub-groups, many scientists were placed under huge strain for a very long period. The workload was enormous and unrelenting. Managing this workload was critical. The SPI-M-O secretariat played a key role in ensuring that the questions posed to SPI-M-O were important, that trivial requests were filtered out and that the questions passed to SMPI-M-O were answerable. This was aided to a large extent by the expertise within the SPI-M-O secretariat. The secretariat consisted of members of the DHSC Health Protection Analytical Team, who had experience of mathematical modelling and pandemic planning. I cannot emphasise enough how important it was to have this technical expertise within the SPI-M-O secretariat. This not only helped ensure that the resources available to SPI-M-O (i.e. the modelling teams) were well used, but also helped condense difficult technical discussions into understandable consensus statements.

6.19. There were issues with resources for SAGE and its sub-groups, particularly at the beginning and towards the end of the pandemic. In January 2020 members of the LSHTM CMMID set up our own working group. In doing this we shifted a large fraction of our human resources to the pandemic effort. At this point there was, however, no specific funding for these activities as most scientists work on short-term grants. We decided to do it anyway and work out funding mechanisms later. The UKRI put in place a grant scheme that was easy and quick to apply for and promised rapid decisions. Other bodies also put similar schemes in place. Our work was mostly funded by an EU Horizon 2020 grant ('The EU grant') and an award from the UKRI scheme. The latter was, however, only for 18 months, as this was the maximum allowed under the award, and the acute phase of the pandemic lasted for over two years. There was no way to extend this UKRI grant, leading to severe staff shortages during the latter phases of the epidemic, specifically during the Delta and Omicron waves. Work on the Delta and Omicron waves was largely funded by the EU grant, which ran for 3 years. In addition, there was no possibility of recruiting and training individuals to take on this extra work. We had to switch experienced researchers from the scientific projects that they were doing onto pandemic-related activities. This meant that the grants that they were employed on had to be stopped while they were moved onto COVID work. Most funders were amenable to this. However, the majority awarded a 1 year no-cost extension to the grant. This meant that staff had to transition back to their original grants in the second year of the pandemic when the Alpha and then the Delta waves

hit. This meant that human resources were stretched extremely thinly during the second and third years of the pandemic and added to the exhaustion of staff. Better, and longer-term funding arrangements need to be made in the future. Longer term flexible funding (for example 5- year renewable consortium grants) need to be made available for pandemic planning and analysis. This would ensure that we have a cadre of staff that are already working on methods and analyses that can rapidly pivot to emergency work when required. It would also give long-term security and flexibility when emergencies do occur.

6.20. SAGE effectively harnessed the power of UK Science to help inform government decision-making during the epidemic. This was an enormous achievement. Previous emergencies on which SAGE had worked were much more limited in terms of scope and size. SAGE had to adapt its ways of working to the scale of the pandemic. It did this by the late Spring of 2020. The switch to online meetings allowed wider participation. Perhaps this should have been adopted earlier. It is now hard to conceive that in-person working will occur again. The scale of the problem also necessitated new subgroups to be formed, studies to be commissioned, such as the National Core Studies **{JE/58 - INQ000237294}**, and mechanisms to be in place to ensure that the latest scientific evidence could be collated, analysed and presented to decision-makers. It is, perhaps, inevitable that this took some time. At the outset of the epidemic SAGE had to rely on the existing committees (NERVTAG and SPI-M) that were always intended to be used during a pandemic and were kept active, partly for this purpose. Having all of the subcommittees in place before the pandemic would have been unnecessary and wasteful, but perhaps going forward we should plan to have some more key committees already in existence and working on pandemic related matters. Testing how they may feed into decision-making in an emergency through the use of exercises would also be important.

7. Infectious Disease Modelling: An Overview

7.1. There is an enormous range of techniques that are employed to understand the spread of infectious diseases and assess the potential impact of different policy options. What follows is a very brief summary:

7.2. Models are mathematical descriptions of real-world systems or processes. They are simplifications of complex systems, concentrating on the key drivers or features. A distinction is sometimes made between statistical (or empirical) and mathematical (or mechanistic) models. The former approach is data-driven and seeks to find patterns and associations. It is the typical approach used in many aspects of the natural or social sciences. It includes descriptive analyses (describing patterns in the data through key summary statistics, such as averages, ranges, measures of the variance, skewness etc), correlation and regression analyses (machine learning methods are variations on regression analyses). In regression analyses the link between the outcome of interest (dependent variable) and explanatory variables can take many different forms, though a limited number of functions are used in practice. Mathematical (or mechanistic) models, on the other hand, are typically built from a theoretical foundation – the links between different variables in the system are assumed. Models typically consist of a number of different elements: variables, which describe some aspect of the system that may change over time (e.g. cases); equations that describe how the different variables might depend on each other and might change; parameters (typically constants) that have a certain numerical value; and starting (or initial) conditions that describe the state of the system at the beginning of the simulation. Such models are often run over time (particularly in epidemiology) and the emergent behaviour of the system is explored. In practice, the distinction between these different approaches (statistical and mathematical models) is often small as parameter values assigned within mechanistic models are derived from statistical analyses and/or estimated through fitting the models to data, using similar approaches used to fit "statistical models". Both statistical and mathematical models are used in outbreak analysis.

7.3. Models can be further subdivided according to whether stochasticity (random variation) is assumed to play a role. Stochasticity can act on parameters (the parameter values are drawn probabilistically from statistical distributions) or processes (the change in a variable is determined by stochastic process), or both. Running a stochastic model many times will generate a distribution of results. Deterministic models, on the other hand, do not take account of random variation. Re-running a deterministic model with the same parameter values and starting conditions will always give the same result. Other things being equal, deterministic models are simpler to analyse and quicker to

solve and fit to data (reducing the computational burden). They can be used to approximate the average behaviour of a system. However, this approximation only holds when the population is large. In practical epidemiology, deterministic models may be used for modelling a large-scale community epidemic. However, stochastic models would typically be required at the outset of an epidemic (when chance may play an important role as cases are few) or at the end of an epidemic (for similar reasons) or when modelling an outbreak in a small setting, such as household, or hospital ward, or care home.

7.4. A further distinction can be made between individual (or agent-based) models, in which each agent in the system is explicitly modelled and tracked over time, and compartmental models in which aggregates of individuals are modelled. A typical compartmental epidemic model might take account of the number of susceptible, infectious and recovered individuals in the population at a given point in time. Indeed, a typical compartmental model used for public health policy, would usually further subdivide the groups in the model to take account of demographic, social or spatial groups as well - allowing a degree of heterogeneity in infection or disease risk to be included in the model. An individual-based model, on the other hand, would explicitly represent every individual in the population (instead of groups of individuals). They are typically more flexible than compartmental models and almost always stochastic in nature (compartmental models can be either stochastic or deterministic). As the status of each individual is explicitly tracked, such models are usually significantly more computationally intensive than compartmental models. This means that they are usually much more difficult to fit to observed data, requiring far greater computing power and perhaps specialist techniques. The added flexibility that they allow typically requires more parameters to be estimated and/or more assumptions to be made. That is, the additional flexibility that arises from an individual-based model comes at a cost to the analyst in terms of data requirements as well as ability to rigorously fit the model to data and speed of this.

7.5. When using mathematical and statistical models in public health a key distinction should be made between forecasting (what we think will happen) and scenario analyses (what might happen given a set of conditions, often including policy options). An analogy with the weather and climate modelling may help make this distinction more clearly. A weather forecast provides a statement about the likely weather in the

next few days (e.g. there is an 80% chance of rain tomorrow). There is no ability to change the weather – it will either rain or not. Climate models have much longer time horizons and give scenarios for, say, mean global temperatures under different assumptions about greenhouse gas emissions. These are scenarios, which cover different policy options. It is possible to choose a policy that (say) minimises global temperature rises under certain constraints. That is, it is possible to choose a possible trajectory - that is, choose a future - though enormous uncertainty will remain. Forecasts are made over a short time horizon – in epidemics this is typically no more than a few weeks. A wide variety of methods can be employed for short-term forecasting, including both statistical and mechanistic modelling approaches. Throughout the pandemic, SPI-M-O commissioned and received short-term forecasts from several different modelling groups on key outcome measures such as the number of hospital admissions and hospital beds that might be required over the next few weeks and the number of deaths. A wide variety of different statistical and mathematical models were used to generate these forecasts. Each individual forecast was examined and compared with other forecasts and outliers were discussed. This set of individual forecasts was then combined into an ensemble forecast that accounted for the uncertainty inherent in each of the separate forecasts. The whole process was repeated every week, with new data being incorporated into updated forecasts. The aim was to give an up-to-date assessment of the current epidemic situation and a probabilistic statement of what may happen in the coming few weeks.

7.6. Scenario modelling is very different. It attempts to answer the question "What might happen if…". Different scenarios might cover a range of different policy options: what might happen if schools are closed; what might happen if non-essential workplaces are closed; what might happen if both schools and workplaces are closed? These scenarios are usually run over longer periods of time than forecasts so the effect of the different scenarios can be ascertained. However, as the models are simplifications of very complex real-world systems, it is highly unlikely that any one scenario will ever give an accurate quantitative forecast over the time period that they cover. That is, these models provide a rough indication of the outcomes that might be expected under different scenarios and can be used to improve our understanding of the complex real-world system and the impact of different policies. Quantifying the uncertainty in projections and presenting this uncertainty is critical. Mechanistic models are generally

used in epidemic scenario analyses as they are designed to take account of the main epidemiological drivers.

7.7. There was a misunderstanding throughout the pandemic of the difference between forecasting and scenario modelling. This was propagated within some areas of the media and the political class. Scenarios were treated as forecasts, and when a particular scenario did not come to pass (usually as deliberate action had been taken to avoid this) then this was treated as a failure of the modelling community or modelling approach. The opposite was the truth. The use of models allows different possible future epidemic trajectories to be explored, giving decision-makers crucial information so that they can choose which policy option is preferred. For instance, both Imperial and LSHTM models suggested that if no action were taken and people continued to behave as they had before the pandemic then hundreds of thousands of COVID-19 deaths would have been expected over the spring and summer of 2020 as the epidemic would pass through the population of the UK infecting the majority of individuals. This was a scenario, not a forecast. This awful eventuality did not take place because the Government took action to avoid it – i.e. they mandated the first national lockdown.

7.8. Variants of the SIR (Susceptible-Infectious-Recovered) structure are at the heart of most mechanistic epidemic models used for policy evaluation. Variants used for COVID-19 typically include an Exposed but not yet infectious class, as well as an Asymptomatically infectious class (see, for instance, **{JE/59 - INQ000212187}**). At their heart these models assume that individuals contact each other at a given rate. This contact rate might vary by social groups to account for heterogeneity in contacts (and therefore risk of infection). In many models of close-contact infectious diseases (such as COVID-19) the contact rates vary by age group as school children have higher average rates of contacts than adults and different age groups typically mix preferentially with themselves (e.g. children with other children at school, and adults with other adults at work). This means that an n*n matrix of effective contact rates must be assumed to reflect this heterogeneous contact patterns with respect to age, where n is the number of age groups in the model. Contact surveys, such as POLYMOD **{JE/45 - INQ000092682}** and CoMix **{JE/60 - INQ000212188}** can be used to estimate the rate at which individuals contact each other within and between age groups. New infections occur when contacts are made between infectious and susceptible

individuals. Thus, other things being equal, the more infectious individuals there are in the model population, the higher the rate at which susceptibles will be infected. As the epidemic proceeds, the number of susceptibles in the population is depleted. Eventually, there are insufficient susceptibles in the population for chains of transmission to be maintained (at this point, each infectious individual produces less than one other infection on average, that is the effective reproduction number is below one). This is the point that herd immunity is achieved. If the level of population immunity is above this level, then infections will decline in the population, without the need for further interventions (such as social distance measures). This state can also be achieved by vaccination, which also typically reduces susceptibility in the population.

7.9. At the outset of a pandemic when the number of susceptible individuals in the population is very large (essentially the whole population were susceptible to SARS-CoV-2) then the depletion of susceptibles can be ignored over the short-term. Branching process models (that do not usually take account of changes in susceptibility) are sometimes used at this point in the epidemic (i.e. at the outset). See for example, **{JE/11 - INQ000092645; JE11A – INQ000255406}.**

7.10. At the beginning of an outbreak of a novel pathogen, there is likely to be very sparse data relating to it (almost by definition). Much early work by modelling teams is devoted to characterising the new pathogen, in terms of key quantities such as the reproduction number (the average number of further cases generated by a typical case), the serial interval (the time between onset of disease in a case and onset of disease in the secondary cases arising from this case), delay distributions (the delay between infection and infectiousness, the delay between disease onset and hospitalisation and death, etc), as well as assessing key indicators of severity, such as the case-hospitalisation ratio or the case-fatality ratio (more reliable measures, as they are not influenced by reporting efficiency, are infection to hospitalisation and infection to death ratios, though these are more difficult to estimate). These early efforts to characterise the pathogen involve analysis of data from the initial focus of the epidemic - often overseas, with the inherent difficulties that this entails in terms of access to high-quality data. Nevertheless, these initial estimates are critical in their own right, and also essential to help build and parameterise (i.e. assign numerical values to parameters in) transmission models. These, in turn, can be used to simulate the spread of the infection in the setting of interest (e.g. the UK). Early simulations from such models

are typically heavily reliant on prior assumptions and the characterisation and parameterisation from overseas data (or even previous epidemics of similar pathogens). Projections performed at this stage are therefore particularly uncertain, as there are few data from the setting of interest to compare the model results to. As the epidemic progresses in the setting of interest, more local data accumulate. Model projections are then compared to these data, by calculating "a goodness of fit statistic" that measures how well the model fits the data. By trying different combinations of parameter values and measuring how well these different combinations fit the data (using the goodness of fit statistic), it is possible to narrow down uncertainty in parameters (and indeed model structure). This is the process of model calibration. By repeating this process as more data become available, it should be possible to further reduce uncertainty in model parameters and structure. Thus, the models are repeatedly fitted (or calibrated) to the emerging epidemiological data. Throughout the epidemic the models used to estimate R, and those used to produce short term forecasts and projections were recalibrated to data every week. This repeated re-fitting of the model to the data means that as more data accumulate, the model projections become more informed by the local data, so the influence of initial assumptions on the results becomes less marked. That is, the findings start to become more data-driven rather than assumption-led. Models can be compared to many different data streams, including cases, hospitalisations and deaths, as well as data on the cumulative fraction of the population who have serological evidence of exposure to the virus (i.e. are antibody positive). In general, the more data that are available to fit the model to and the more different types of data are available, the better.

7.11. Uncertainty will always remain, and multiple assumptions are inevitably made, particularly at the early stages of an emerging epidemic. Sensitivity analysis is a means by which the sensitivity of key results to changes in model parameters or assumptions is tested. It is therefore critical for assessing the robustness of policy conclusions. Sensitivity analysis can take many forms: the simplest and most common type of sensitivity analysis is a one-way sensitivity analysis, in which a single parameter or assumption is changed, and results are outputted for different values of this parameter (different assumptions). It is often the case that model results are sensitive to small changes in some parameters, but not very sensitive to changes in other parameters. In this way research or surveillance priorities can be identified, allowing targeting of

efforts to reduce uncertainty in the most influential parameters. Multivariate sensitivity analysis involves changing multiple variables simultaneously. This is usually done by drawing combinations of parameter values probabilistically from input distributions. The resulting model output can be examined statistically to see how changes in parameters, or combinations of parameters, affects results. There are close links between this sort of sensitivity analysis and the process of model fitting (calibration).

7.12. Each data stream that a model is fitted to is always affected to some extent by biases and delays. For example, not every infection results in a clinically apparent case (i.e. some infections are asymptomatic), and not all of these will be tested and reported (that is, the case data are biased downwards). In addition, the testing and reporting of cases introduces delays. For COVID-19 the delay between infection and symptoms (the 'incubation period') was about 5-6 days initially, but with large variation between individuals. Added to that is the delay from symptom onset to test (if testing occurred at all) and a further delay for the results to be available. In the early stages of the pandemic in the UK this delay to confirmation was about another 6 days, again with large variation between individuals. So, cases being reported on a given day actually reflected infections that may have occurred 10-12 days earlier. Hospitalisation typically occurred about a week after symptom onset, and death another week or so later. So, each of the data streams have different delays attached to them as well as different biases. Understanding and measuring these delays is critical to understanding the current and short-term future epidemiological situation. In early March 2020 the UK started reporting daily numbers of cases. At this point these numbers were relatively small (dozens of cases). However, these numbers did not reflect the number of infections that occurred that day, but the number that had occurred, been detected by our surveillance system, and tested positive some time earlier (roughly 12 days earlier, on average). As the epidemic was growing very quickly at this time (analyses suggested a doubling time of about 3 days) the actual number of infections that occurred on a given day would have been far greater than was being reported (the epidemic would have been expected to double 4 times over that 12-day delay, meaning that the size of it would be roughly 16 times bigger than was being reported. If not all cases were being identified and tested – which was the case – then the real size of the epidemic would be even bigger). Likewise, when an intervention is put in place to reduce transmission (e.g. lockdown) then there would be a delay to seeing an

epidemiological effect. Reported cases would continue to increase for some time (again, about 12 days using the assumptions above) because infections that have already occurred will take time to become apparent, be tested and then reported. The lag to a decline in hospitalisations and deaths would be even longer, so it would be expected that cases start declining while hospitalisations and deaths continue to increase. By analysing these delays (and how they might be changing over time) and incorporating them into models, it is possible to give a more accurate picture of the current epidemiological situation than can be derived from the raw surveillance data. That is, models combined with careful epidemiological analyses can help correct biases and delays in reporting to greatly improve situational awareness.

7.13. The process of fitting models to data is sometimes referred to as calibration (see earlier). Validation involves checking the validity of the model. This might be done by comparing model simulation results against an external data set – one that was not used in the model construction or parameterisation. Examples might include comparing the fraction ever infected in the model with the fraction who have serological evidence of infection (antibodies) – assuming that these data were not used in the calibration process – or comparing model projections with data that subsequently emerge. Model projections are impossible to validate in real time (as the events have not happened yet). However, projections from independent models can be compared with each other, to assess whether there is broad agreement between them and examine the reasons for any differences. This form of cross-validation was adopted by SPI-M-O throughout the pandemic.

7.14. At the heart of transmission dynamic epidemic models is a set of assumptions governing how individuals contact each other, as transmission occurs during these encounters. Information on these contact patterns can be gleaned from contact surveys, which ask individuals to record who they contacted over a given time-period. Pre-pandemic contact surveys (such as the POLYMOD survey **{JE/45 - INQ000092682}** were critical in determining baseline rates of contact for use in epidemic models. However, these rates of contact were likely to change in a severe epidemic, if individuals take action to try to reduce their risk (i.e. in response to perceived risk) or because governments mandate that they change their behaviour, by imposing restrictions. For this reason, the CoMix survey was undertaken to monitor how these epidemiologically relevant contacts were changing over the course of the

epidemic. This was also supplemented by data on mobility, which was available from other sources, notably Google **{JE/61 - INQ000212189}**. However, these measures are retrospective – i.e. CoMix gave estimates of contact patterns over the previous week. It is still not possible to quantitatively predict how behaviour might change as a result of changing epidemic risk or government action. The lack of any quantitative behavioural projections therefore widens the uncertainty surrounding epidemic projections. Epidemiologists have to rely on assumptions about how behaviour might change (or might not change) and sensitivity analyses to these assumptions (see for instance, **{JE/62 - INQ000212190}**). An improved quantitative understanding of how behaviours might change under different epidemiological circumstances would directly help public-health decision makers and improve epidemic models.

7.15. In summary, there is an enormous range of analyses and models that are typically undertaken during an infectious disease outbreak. Many of these analyses (particularly those related to forecasting and projections) are repeated frequently (often every week), requiring an enormous quantity of resources in terms of manpower, computing and data that needs to be sustained over the course of the epidemic. There are many possible uses for models and analysis over the course of an epidemic, a few of which have been outlined above. However, the three main reasons are:

7.15.1. To increase understanding. The process of constructing and describing a model helps organise thinking and improve transparency of assumptions. Constructing and parameterising models (quantifying parameter values) also helps identify knowledge gaps which can help research prioritisation or the design of surveillance systems.

7.15.2. Forecasting and nowcasting: models are used to make short term predictions of outcomes of interest or to provide improved situational awareness.

7.15.3. Scenario analyses, in which models are used to explore different possible future states of the system (epidemic trajectories) given different assumptions. Scenario analyses can be used to explore the possible impact of different policies to try to mitigate the epidemic.

7.16. Epidemic models appear to be most strongly associated with forecasting in the public consciousness. Whilst models are put to this use during outbreaks (and were during the COVID-19 pandemic), they are probably more commonly used to address the other

two areas: i.e. improving understanding of the underlying processes; and scenario analyses. The confusion between forecasts and scenario analyses should be addressed to help improve public discourse on the use of models in many areas of the natural and social sciences. Models typically use a combination of assumptions and data (that informs model parameter values). If models are continually re-fitted to emerging data (as was done during the COVID-19 pandemic by the groups contributing to SPI-M-O) then, other things being equal, uncertainty in parameters and model structure can be reduced through this process and model results should be increasingly data-driven. Validation (comparing model results to data not used in the calibration process) can be performed to increase confidence in model results. Model projections can only be validated to data retrospectively (as the future has not occurred yet), but comparisons to other model projections can be done in real-time to help understand the reason for any differences and better define the uncertainty in projections. This was done by SPI-M-O throughout the pandemic. Nevertheless, significant uncertainty will always remain, and this should be estimated and presented to decision-makers and the public alike. Sensitivity analyses should be used to check the robustness of results to key remaining uncertainties and to guide future research. One obvious remaining gap in our knowledge is a quantitative understanding of behavioural change in response to epidemic risk, public health messaging, and the imposition of public health measures. There are many others, of course. Research is needed to reduce these knowledge gaps to enhance our ability to respond appropriately to epidemic threats.

8. The Early Stages of the Pandemic

8.1. The national and international scientific effort to describe this new disease, how fatal it was, who was most at risk and what could be done to treat them better, identify, characterise and understand the virus that caused the disease, track its spread and ascertain how it might be contained or controlled was enormous. SAGE's work in the early part of the pandemic (January to March 2020) also focussed on reviewing the data on the clinical, virological, phylogenetic and epidemiological features of the disease (COVID-19) and its causative agent (SARS-CoV-2) as well as assessing the potential behavioural response to the epidemic and measures put in place to control it.

8.2. In January and February 2020, when there were very few cases in the UK, efforts concentrated on international data. Understanding the epidemic in its initial focus of Wuhan, Hubei Province, China was critical. However, this was difficult. Initially, the Chinese authorities reported that all of the cases were linked to a seafood market in Wuhan. Indeed, official case numbers remained static for much of January 2020, suggesting that the virus could have limited ability to spread between humans. Details of the cases (which would allow epidemiological and clinical characterisation) were also difficult to come by and many of us were using an unofficial website that compiled details from trawling local news and social media reports. A key source of information was cases that had been exported from China. By assessing the number of these exported cases and whether they were increasing, some inference could be made about the true state of the epidemic in Wuhan. These early estimates suggested that there were many thousands of unreported cases, rather than the tens of cases reported by China **{JE/63 - INQ000212191; JE/64 - INQ000212192}**. Estimates of the scale of the epidemic in Wuhan could also be made from the prevalence of infection in repatriated nationals. Many countries arranged repatriation flights in late January or early February, including the UK. Other countries (e.g. France and Korea) tested all of those who were repatriated (sometimes on multiple occasions). Initially the UK did not, but on recommendation from SAGE this was instigated. By looking at the prevalence of infection in these individuals and whether it was changing over time (in later repatriation flights) it was possible to get an idea of the scale of the epidemic in Wuhan. By fitting a model to these multiple data sources (reported cases, exported cases, repatriated cases) it was possible to derive a more accurate estimate of the size of the epidemic in Wuhan, the level of under-reporting and, by mid-February, an early estimate of the impact of the lockdown **{JE/06 - INQ000092695; JE/06A – INQ000255394}.**

8.3. By late January, Chinese researchers were beginning to publish the results of their initial analyses. These papers gave crucial epidemiological and clinical information, allowing key parameters to start to be estimated, such as the serial interval (the time between generations of cases), the reproduction number (the average number of further cases each case generates) and the incubation (time from onset to symptoms) and infectious periods. In addition, data on severity also started to accumulate, but as virtually all cases were hospitalised for infection control reasons it was difficult to

estimate what fraction of cases needed to be hospitalised for clinical purposes, and due to delays between infection and death, estimates of the case-fatality ratio (the fraction of cases that die) had to be made very carefully, to account for potential biases (see below), and took longer to ascertain. Nevertheless, intensive efforts to try to better quantify these critical parameters continued throughout the early part of the epidemic.

8.4. Establishing parameters such as the incubation period and infectious period distributions, as well as case definitions, are not only critical for developing infectious disease models and understanding the epidemiology but are essential for designing public health control measures as they determine who needs to be investigated further (and potentially isolated), how long a case needs to isolate for, and how long a contact must be in quarantine. The email exchange instigated by Chris Whitty on 9th March 2020 is on this issue where he asks whether isolation should be for 7 or 14 days **{JE/65 - INQ000212193}**. Interestingly, I clearly make a mistake in this exchange (confusing the time since infection with the time since onset), though Professor Neil Ferguson corrects it – indeed, I always thought that 14 days isolation was longer than necessary **{JE/66 - INQ000212194}**. The case definition determines who these measures would apply to. In early March 2020 there was discussion of a change to the case definition used in the UK (NERVTAG 13th March 2020 **{JE/67 - INQ000212195}**) to one that considered just fever and cough. By that time data were available from China that followed up cases and their contacts **{JE/68- INQ000212196}** and from the first few hundred (FF100) database in the UK. It was possible to see from these data that about 20% of cases would be missed if such a definition was adopted but including muscle ache and fatigue increased the sensitivity to about 87% **{JE/69 - INQ000212197; JE/70 - INQ000212198}**.

8.5. When interventions were put in place in China in late January 2020, much effort was directed to trying to understand the impact of these measures (which were generally very strict). Due to delays between implementing such measures and their effect being apparent in the epidemiological data (as there is a delay between infection and clinical signs and further delays to testing and reporting results), as well as difficulties inherent in understanding the Chinese data, it took until mid-February before the impact of these policies could start to be quantified.

8.6. One of the difficulties of the Chinese data was that it was not clear what the "case definition" was – i.e. how cases were being defined, and therefore counted. A plausible

explanation for the apparently flat epidemic curve in mid-January 2020, and that all cases were linked to the seafood market in Wuhan, was that this was part of the case definition (i.e. individuals with a link to the seafood market were investigated and tested, but others were not). That is, those with symptoms who had no link to the seafood market were not investigated and could therefore not count as cases. As the market had closed, this would account for the apparently flat number of cases, the sudden surge in cases at the end of January being a result of relaxing this definition so that cases did not have to have a link to this market. A similar situation may also have occurred for cases elsewhere (i.e. if part of the case definition (which would govern whether someone was tested or not) was that a person with symptoms also had to have travelled from Wuhan, or be a contact of someone who was from Wuhan, then again the epidemic would look flat in other Chinese cities, without necessarily being so. Members of SAGE used their personal connections with Chinese researchers to try to clarify issues around their data and there were attempts to reach out through official channels **{JE/71 - INQ000212199; JE/72 - INQ000212200; JE/73 - INQ000212201; JE/74 - INQ000212202}**. However, doubts remained about the validity of some of the data and the lack of clear case definitions from the Chinese authorities hampered its interpretation. Indeed, later analysis by a team from Hong Kong confirmed our early suspicions about the Chinese case definitions: for most of January, cases were only counted if they had visited the Huanan Seafood Market in Wuhan; and cases elsewhere in China were only counted if the case had visited Wuhan. Better clarity and openness from the Chinese authorities or - failing that - better contact with Chinese researchers and clinicians on the ground could have improved our understanding of the clinical and epidemiological features of this new disease and the impact of the measures that had been put in place during these critical early weeks.

8.7. Estimation of the fraction of cases that die (the case-fatality ratio) is fraught with difficulties when done in real-time. First, it is difficult to ascertain how many cases there have been, as not all cases are tested and reported – that is there is an unknown level of under-reporting, which will likely differ by setting. A better measure is therefore the infection-fatality ratio, but this requires data on the number of people who have been infected, which is usually not available (blood tests can be done on a defined population to assess how many have been infected, but this is difficult and time consuming and the necessary tests take time to develop for a novel pathogen). Hence,

although infection fatality rates are theoretically better, as they are not dependent on the efficiency of the surveillance system, they are rarely available in practice. Secondly, there is a delay between becoming a case and death – for COVID this delay is typically of the order of 2 weeks but is highly variable across individuals. Thus, counting the deaths and dividing by the number of cases will give a highly biased estimate of the fraction who are likely to die if the epidemic is growing rapidly, since the appropriate denominator is the fraction of cases that occurred two weeks earlier (not cases occurring now). There are a variety of ways to account for this bias, but one of the simplest methods is to divide the number of deaths by the number of resolved cases (i.e. they have either died or recovered). This should give an unbiased estimate of this critical parameter – as I explained to NERVTAG on 21st January 2020 **{JE/75 - INQ000023119}**. The delays between infection and deaths also inevitably mean that there are lags to get accurate estimates of these case or infection fatality ratios.

8.8. Early SAGE meetings (January and February 2020) spent considerable time assessing what the implications of this new disease might have if / when the epidemic came to the UK. This was done through comparison of the possible epidemic trajectory of COVID-19 to the "Reasonable Worst Case" planning scenario. This "Reasonable Worst Case" scenario was based on the 1918 pandemic of influenza and consisted of assumptions about the speed of spread, the size of the resulting epidemic wave in terms of numbers of people infected, clinically ill, fraction hospitalised and the number of deaths. It is supposed to represent the worst-case scenario that might reasonably be expected, and largely represents an unmitigated severe influenza pandemic. Much time in the early SAGE meetings was devoted to comparing the possible epidemiological pattern from an unmitigated COVID-19 pandemic in the UK to this Reasonable Worst Case. Whilst there was significant uncertainty and it was clear that there were differences - the longer serial interval meant that the disease would spread more slowly than influenza, for instance - the early indications were that an unmitigated epidemic of COVID-19 would be qualitatively similar to the Reasonable Worst-Case planning assumption. As the reproduction number was estimated to be between 2 and 3 and the whole population was susceptible then an unmitigated epidemic could affect a large fraction of the population if people did not change their behaviour to effectively reduce their risk. Given the indications that the infection fatality ratio was about 1% (estimates varied between about 0.5% and 1.5% by mid-late February 2020) then such

a large wave of infections would be expected to lead to hundreds of thousands of deaths. Estimation of the impact on hospital bed-usage was more difficult, as in China and elsewhere cases were hospitalised irrespective of symptoms to prevent community transmission. In most instances, they had to remain in hospital until they were deemed no longer infectious to others. In practice this usually meant until they had tested negative on two consecutive PCR tests. Thus, the fraction of cases that needed to be hospitalised on clinical grounds and the average length of stay in hospital was not readily available from the early, overseas data. Nevertheless, it was obvious that a large unmitigated wave would be likely to put the NHS under extreme pressure. Reliable quantification of the likely hospitalisation rate and length of stay came quite late – the meeting on Sunday 1st March between Imperial College, LSHTM, University of Oxford and NHS teams was helpful in this regard (this is referred to later in this statement).

8.9. SAGE also considered the possible impact of interventions that could be put in place during the early phase of the pandemic, to delay and reduce its impact. The two key early measures were travel restrictions and contact tracing. By mid-February we were also considering the possible impact of community interventions. The most obvious was school closures, given the reasonably good evidence base on the effectiveness of this intervention at reducing influenza incidence (reviewed here: **{JE/76 - INQ000212204}**). Note that these interventions; travel restrictions, contact tracing and school closures were the subject of an email trail instigated by Prof Chris Whitty on 29th January 2020 as he was seeking information on the effectiveness of potential measures **{JE/66 - INQ000212194}**. Other interventions were also looked at, including individual and household-targeted interventions as well as measures aimed at slowing community transmission. Given the difficulties in accurately quantifying the clinical picture associated with COVID-19 infection (particularly surrounding hospitalisation) and the time taken to develop and parameterise mathematical models, little detailed modelling work was done on the impact of these social distance and community measures in the UK until late February and early March 2020. I cover the work that LSHTM undertook on the potential impact of non-pharmaceutical interventions later in this statement.

8.10. There are a range of measures that can be put in place to reduce the likelihood of infectious travellers entering the country of interest (in this case the UK). These include

giving travel advice, entry and/or exit screening (which can be based on temperature and/or symptoms, or tests), travel restrictions and border closures. Before the COVID-19 pandemic the evidence base related to these measures was mainly concerned with influenza. These studies were reviewed by the WHO in 2019 **{JE/76 - INQ000212204}** which concluded that they were generally inefficient at finding cases and resulted in relatively small delays in the epidemic (the benefits being measured in days). Work that I had been involved in 15 years earlier looking at the potential impact of travel restrictions to delay the spread of pandemic influenza **{JE/77 - INQ000212205}** was part of this evidence base. It suggested that unless travel restrictions were extremely stringent, they were only likely to delay a pandemic by a few days or weeks. Given the likely low efficacy of these measures WHO did not recommend entry or exit screening or border closure to control pandemic influenza **{JE/76 - INQ000212204}**. At the time of the email query from Professor Chris Whitty on the 2nd of February 2020, similar arguments could be made about the effectiveness of travel restrictions on passengers from China: they were likely to buy little time, unless they were very stringent and coordinated internationally **{JE/78 - INQ000212206; JE/79 - INQ000212207}**. Nevertheless, these measures were assessed by SPI-M, NERVTAG and SAGE at the outset of the epidemic (the paper on travel restrictions for influenza was considered by SAGE on 3rd February 2020 **{JE/80 - INQ000212208}**). In addition, we developed a model to look at the potential impact of travel restrictions at delaying COVID-19 **{JE/07 - INQ000092696; JE07A – INQ000255395; JE07B – INQ000255396; JE07C – INQ000255397; JE07D – INQ000255398; JE/81 - INQ000212209}**. Overall, as with influenza, the work suggested that travel restrictions would likely result in relatively small delays to the epidemic, unless they were very stringent (border closure). As border measures do nothing to slow transmission within a country they must be adopted very early to be effective (before transmission is established within the country) and must be maintained until the end of the international epidemic or the country is effectively immunised.

8.11. The 2019 WHO advice on NPIs for influenza recognised the poor standard of evidence available in relation to the effectiveness of contact tracing and isolation. However, they recommended home isolation of symptomatic individuals, but did not recommend home quarantine of exposed individuals or contact tracing to control a pandemic of influenza **{JE/76 - INQ000212204}**. The reasons for this are multiple but include the

very rapid serial interval (time between generation of cases) for influenza (usually of the order of 2-3 days) and the risk that individuals can become infectious to others before they become symptomatic. In addition, contact tracing is extremely labour-intensive. When an epidemic is growing quickly, the resources devoted to contact tracing can become overwhelmed, rendering the system of little use to slow or prevent transmission. The early Chinese investigations (published in late January and early February 2020) suggested that COVID-19 appeared to have a longer serial interval and incubation period (time from onset to symptoms), which may make contact tracing and isolation a more feasible strategy than it would be for influenza. For these reasons it was implemented as a means of slowing or preventing transmission in the early phase. Modelling studies were instigated to look at the potential effectiveness and were considered by SPI-M-O and SAGE. I cover CMMID's early work on this later.

8.12. Towards the end of February and early March 2020 SPI-M-O were repeatedly asked about the potential impact of closure of sporting events. It was also discussed at SAGE. Our models are not so detailed that they distinguish attendance at sporting events – far from it – so it was not a question that could be easily addressed from a modelling point of view, and there was (to my knowledge) no data on the effectiveness of these measures. Although sporting events might involve many thousands of individuals, they are generally held outdoors (indoor events would be higher risk) and close contact is probably only made with a relatively small number of different individuals (those seated close by, drinks vendors etc). Hence, the events themselves were not thought to be a high risk and stopping them was unlikely to result in a significant impact on epidemic progress, though reducing leisure-related contacts more generally would have a bigger effect **{JE/82 - INQ000212210}**. Stopping attendance but allowing pubs to stay open (where people might gather to watch games) could potentially exacerbate the risk. Although there was no data to inform these assumptions at the time, the Events Research Programme did address the risk associated with attending such events in the summer of 2021. A variety of methods were used, including environmental and behavioural monitoring at events and the conduct of a large- scale self-controlled case-series analysis to assess the risk of attendance. The CO_2 monitoring suggested that air quality was generally good at all venues and the self-controlled case series found no evidence of an increased risk from attendance at a wide variety of sporting and

cultural events except multi-day festivals, though some control measures were in place at the time **{JE/83 - INQ000212211}**.

8.13. As cases started to accumulate in the UK, significant attention turned to the estimation of the growth rate – often expressed as the doubling time (i.e. how long it takes for the number of cases to double). This was extremely difficult to estimate accurately, due to our very poor surveillance and data-handling systems at the time. Before the massive expansion of testing in April 2020 (during the first national lockdown) access to tests was restricted to certain patient groups – largely travellers with symptoms consistent with COVID-19 who had returned from an affected country or area (e.g. China), or the contacts of known cases. This meant that case numbers being reported in the UK were potentially more reflective of epidemics overseas than the epidemic here. To try to get a relatively unbiased estimate of the prevalence of infection here and whether it was changing over time, two surveillance systems were put in place. First, all individuals in intensive care units (ICU) with pneumonia were to be tested. Second, the Royal College of General Practitioners (RCGP) Sentinel Surveillance system were to test a random sample of patients attending GP practices for acute respiratory illness for SARS-CoV-2. All tests at the time were PCR tests, which would build in a delay to the results being available. These two surveillance systems were launched in late February 2020. Positive cases were found through these systems almost immediately. This was deeply worrying, as the sensitivity of the ICU system had been estimated by Professor Neil Ferguson to be about 1 in 1000, and the sensitivity of the RCGP system was estimated to be about 1 in 1500. That is, any cases detected through these systems might indicate a large underlying epidemic. Furthermore, the change in cases reported through these systems should more accurately reflect changes in the underlying epidemic in the UK than the PHE contact-tracing-based cases. I first started receiving and analysing these data around the 8th of March 2020. Unfortunately, they were very difficult to analyse. First the delays between onset in cases to confirmation and entry onto the database were long at about 5-7 days but with a very long tail so that some cases were taking up to 3 weeks to be recorded. That is, a case would only be registered about a week after it had first shown symptoms (often much longer) and because of the delay between infection and first symptoms (typically 5 or 6 days) new cases registered on a given day were actually reflecting infections that might have occurred two weeks earlier (I brought these significant delays to the attention of SAGE

on 13th March 2020, see **{JE/84 - INQ000212212}**. With the epidemic doubling every few days, this could mean that case numbers being reported on a given day (with the impression that they had been infected on that day) could be out by orders of magnitude. Also, estimating the growth rate of the underlying epidemic (as reflected in the RCGP and ICU systems) was made more difficult by not distinguishing these cases in the database (this database is often referred to as the "line-list" as each line on the database represents an individual patient). Instead, cases from these systems were simply mixed in with the contact traced cases. After discussion with PHE, the "sporadic" cases (i.e. those picked up through the RCGP or ICU systems) were marked as such, but it was still not possible to tell which system they had come through. Given these systems had different sensitivities and different delays (it typically takes a week after onset to be hospitalised) it was important to be able to distinguish them. We never resolved this issue properly. Despite all these caveats and uncertainties, analyses of these data revealed that there were significant delays to confirmation and reporting of cases and that the scale of the epidemic was far larger than was being reported. Both of these had serious implications for the accuracy of the information being portrayed to decision-makers and the public (and even other SAGE members) at the time. In addition, it was very difficult to determine the growth rate. Analysis of the crude data (including all cases irrespective of why they were being tested) suggested that the doubling time in the UK was around 3 days. However, analysis of the "sporadic" cases suggested a longer doubling time (slower epidemic) of 5-7 days, although these estimates came with very wide confidence limits. Email exchanges and papers that cover these methods during mid-March are given in **{JE/82 - INQ000212210; JE/85 - INQ000212213; JE/86 - INQ000212214; JE/87 - INQ000212215; JE/88 - INQ000212216; JE/89 - INQ000212217; JE/90 – INQ000212219; JE/91 - INQ000212040}**. The net effect was, however, that in mid-March, when tens of new cases were being reported on a daily basis, we were estimating that there were hundreds or perhaps thousands of cases occurring every day. Implementing radical measures to limit transmission was probably one of the most challenging decisions any government has faced since the Second World War. It was made much more difficult by having very poor data. I don't know whether decisions were hampered and/or delayed because of this very poor situational awareness but suspect that it may have been a contributing factor. The ramping up of testing in April 2020 and the establishment of the National Core studies, particularly the ONS Coronavirus Infection

Study, transformed our ability to track the epidemic. The difference in surveillance was like night and day.

8.14. In addition to the growth rate, the reproduction number (or R number) is also a critical measure, as it gives the average number of secondary cases a typical case will generate. If this is greater than 1, then an increase in cases would be expected (that is the growth rate of the epidemic would be positive). It also gives an indication of how effective non-pharmaceutical interventions must be to stop epidemic growth. A simplified example will help to illustrate the point: if the reproduction number is two and interventions aimed at reducing contact are put in place, then contacts need to be halved (assuming random mixing in the population) to stop the epidemic from growing. Measuring contact patterns and the up-take of other interventions along with calculating the reproduction number over time, can therefore give critical information to decision-makers. In addition, an estimate of the basic reproduction number (the reproduction number when everyone is susceptible, and no interventions are in place) is also very important as it gives an indication of the maximum reproduction number for the pathogen in the population. It also gives an indication of the level of population immunity necessary to stop the epidemic without further measures. At the outset of the epidemic, the reproduction number was estimated from settings where cases were spreading (mostly China). The team at LSHTM undertook to systematically review reproduction number estimates from the international literature to enable our model to be parameterised. Early estimates of the basic reproduction number (i.e. without interventions) were uncertain, and ranged from about 2 to about 4, with the most likely estimates being between 2.5 and 3, see Figure S2 in **{JE/59 - INQ000212187}**. SPI-M started providing consensus estimates of the reproduction number from February 2020, largely based on overseas analyses. As the epidemic became established in the UK, the reproduction number could be estimated locally. This started to occur in March 2020. Given the importance of the reproduction number and growth rate estimates for tracking the impact of interventions here in the UK, weekly consensus estimates were estimated for the UK as a whole as well as by region and constituent UK country. However, these weekly, systematic consensus estimates took time to establish – only really coming on stream in April and May 2020 when systems were in place for the data to be shared with groups and processes and statistical methods established for

individual estimates to be shared with SPI-M-O and combined into consensus estimates.

8.15. Similarly short and medium-term projections also came on stream after the first lockdown, as data access was in place for all groups, data streams became routinised and methods were developed in the individual groups to produce these estimates, and at the central level to analyse the individual projections and combine them into a statistically defensible ensemble estimate. The team at DSTL took the lead in generating these ensemble estimates of projections, growth rates and reproduction number estimates over the course of the epidemic.

8.16. I have described, above, some of the evidence that SAGE and its subgroups were considering during the early part of the pandemic (from January to mid-March 2020). I describe, below, some of the specific work done by CMMID to contribute to this knowledge base. Our work also fell into the same categories of: characterising the new disease, its severity and transmission patterns, estimating the potential impact of early interventions to delay cases in the UK (travel restrictions and contact tracing), estimating the impact of other NPIs aimed at slowing or halting the spread of the epidemic once it had become established here, and forecasting the requirements for hospital beds and other resources. I give a brief outline of some of this work below:

8.17. It was critical to understand what fraction of infections or clinical cases die and how this changed by age group. The Diamond Princess outbreak was relatively large and well-studied with a clear denominator (the passengers and crew on the ship). In mid-February 2020 we used the data from this outbreak to estimate the overall infection fatality ratio, **('IFR'),** (which we estimated to be just less than 1%), though this masked large changes with age, with the IFR being estimated to be over 7% in the over 70 year olds **{JE/13 - INQ000092647}**.

8.18. By the middle of March 2020, there were early indications that immunity from natural infection with COVID-19 might not last for very long (see for instance, NERVTAG minutes 13th March 2020 **{JE/67 - INQ000212195}**). There had been a case report of a reinfection from Japan and analyses of other coronaviruses showed that immunity could be short-lived. This would require changing the model structures that most groups were adopting to allow for waning immunity and repeated reinfection. In practice, however, as a lockdown was introduced, the levels of immunity remained

relatively low in the population throughout 2020. That is, waning immunity and reinfection only became a major public health issue after there was widespread immunity in the population (the latter half of 2021). By this time all groups working for SPI-M-O had waning immunity included in their models, though it remained difficult to estimate with any accuracy how long immunity to different aspects of COVID-19 infection (infectiousness, disease, severe disease) might last and how it changed with each new variant.

8.19. Understanding the epidemic data in Wuhan was challenging for many reasons, some of which are outlined above. For this reason, we used the data from Wuhan in conjunction with exported cases and estimates of the level of infection in repatriated individuals to assess the potential scale of the epidemic in Wuhan and the impact of interventions **{JE/92 - INQ000212221}**. Critically, this work showed that, although the epidemic in Wuhan had likely been much larger than was reported, it did not infect a large part of the population, and that the control measures put in place by the Chinese authorities had been effective at reducing transmission (rather than "herd immunity" having been achieved).

8.20. Contact tracing and case isolation is one of the few interventions that can be put in place at the outset of an epidemic to slow spread. Hence, we started to work on a model to assess the potential effectiveness of this approach in January 2020. Professor Chris Whitty had asked about measures that might be used to slow the epidemic on 29th January and I referred to this paper in my reply **{JE/66 - INQ000212194}**. The results of this work were first put online on the 7th of February 2020 **{JE/11 - INQ000092645; JE11A – INQ000255406},** submitted and discussed by SPI-M-O and were published by Lancet Global Health on the 28th of February 2020 **{JE/93 - INQ000212222}**. The model results suggested that contact tracing and isolation could control an emerging outbreak, but a high fraction of contacts (70-80% in most scenarios) would have to be traced and isolated for this approach to have a high chance of success. Increasing transmissibility or increasing pre-symptomatic or asymptomatic transmission rendered control via these methods less feasible. This paper clearly had a major impact in the scientific and public health communities, being one of the most highly cited publications that I have been involved with. It is not clear to what extent it influenced decision-making in the UK, possibly because early data

from PHE indicated that there were significant delays to cases being tested and therefore isolated.

8.21. Detailed estimation of the impact of different non-pharmaceutical interventions in the UK mostly took place in March 2020. The models used by both Imperial College and LSHTM were sufficiently developed by then and data estimates were available for most critical parameters. A critical meeting was held on Sunday 1st March 2020 between the Imperial College, LSHTM, NHS and Oxford University teams that helped establish base-case assumptions for age-specific case-fatality and hospitalisation ratios in addition to the likely average duration of stay in hospital with COVID-19. This meeting was requested by SAGE and was also attended by the SPI-M-O secretariat. The different teams compared methods for estimating infection- and case-fatality ratios and agreed on a range of parameter values (despite different approaches, estimates by the LSHTM and Imperial College teams were similar). The fraction of cases that were likely to require hospitalisation on clinical grounds was established, thanks to some data shared by Prof Peter Horby, as was the likely length of stay in UK hospitals (from analysis of NHS data on viral pneumonia). The agreed parameter values derived at this meeting allowed both teams (Imperial College and LSHTM) to undertake detailed simulation studies on the impact of different NPIs in the UK. The initial results of this work was shared with SAGE in time for the meeting on Tuesday 3rd March **{JE/94 - INQ000212223}**. The paper used our age-structured, stochastic transmission-dynamic model (that had been presented to SPI-M-O at the end of February 2020). Age specific-hospitalisation and infection-fatality ratios were based on the parameters agreed on 1st March, values for the basic reproduction number were drawn from the distribution given in our ongoing review of the literature (so that uncertainty in this parameter was included in our analyses) and contact patterns were based on the POLYMOD study **{JE/45 - INQ000092682}**. We had to make a number of key assumptions about how different NPIs might reduce contacts (this was unknown at the time). The paper suggested that an unmitigated epidemic, in which people did not change their behaviour, would cause hundreds of thousands of deaths and overwhelm the health service, with peak demand for ICU beds being many times greater than the number available. We looked at a range of NPIs including closing schools, shielding the elderly (which we referred to as cocooning), social distance measures, and combinations of the above. The results suggested that although these measures would

likely reduce the burden of disease, they would still likely lead to very large numbers of deaths and demand for health services if implemented alone. Furthermore, a combination of all the measures could be sufficient to supress the epidemic but relaxing this would lead to a resurgence of cases. Overall, although the results were preliminary and highly uncertain, this paper - along with a similar one prepared by Professor Neil Ferguson at Imperial College – gave an early quantitative insight into the scale of the possible epidemic in the UK, as well as the impact of non-pharmaceutical interventions. This initial work was improved over the next few weeks and the model was used to explore an expanded range of different interventions (e.g. the impact of the timing of school closures and the potential impact if grandparents provided care was explored in a paper on the 17th of March **{JE/95 - INQ000212224}**). The various results were provided to SPI-M-O and SAGE as short reports, which we gathered into a summary paper that was released on our repository on the 1st April 2020 and (after peer-review) published by the Lancet Global Health **{JE/96 - INQ000212225; JE/96A – INQ000255451}.**

8.22. Analyses of other policy options came at different points in the epidemic. For example, the epidemiological impact of support bubbles was addressed in the later stages of the first lockdown, when this first emerged as a potential policy. The initial work on the efficacy of contact tracing (including digital contact tracing) was refreshed when it became clearer what the different options were under consideration in the UK (end of April and early May 2020) in the run-up to the launch of NHS Test and Trace. These analyses will be covered elsewhere in this statement.

8.23. As the epidemic became established in the UK a number of related issues became more important, namely, how many infections were occurring, how quickly the epidemic was growing, and what the demand for hospital care would be in the near future. I have outlined, above, the difficulties we faced in trying to establish the size of the epidemic and its growth rate, due to inadequacies in the data. We concentrated our analyses on the "trusted" part of the dataset, which was data that had been collected some weeks earlier, due to the delays to reporting (see earlier). Having established this, we could project forward, using simple fitting of a log-linear model to the daily incidence of symptom onset within the trusted period **{JE/97 - INQ000212226}** to give estimates of bed requirements in the near future. This work went through various iterations at the time and had been discussed by a small group

of PHE analysts on 12th March as well as SPI-M-O the following week. A version was available to SAGE participants on the 23rd March **{JE/98 - INQ000212227; JE/99 - INQ000212228}**. Estimating the delays in the data-streams was important as it helped establish by how much the reported cases were under-represented (see SAGE minutes 13th March **{JE/84 - INQ000212212}**). It also helped establish that critical care bed capacity might be very seriously stretched by the end of March unless measures were rapidly taken to reduce the epidemic.

8.24. After the first national lockdown was instigated on the 23rd of March 2020, there was an immediate need to know how effective it was. The obvious way to measure this was the reproduction number. Just before the lockdown it was estimated to be around 2.6 (that is each case, on average, caused 2.6 other cases). For the lockdown to be effective, the reproduction number would have to be reduced to below one, which would indicate a decline in the epidemic. However, the delays inherent in the surveillance system meant that this would not be apparent for two weeks or more as there was a 5-6 day delay (on average) from infection to onset of symptoms and a similar delay to reporting of the case. Changes in hospitalisations or deaths would occur over an even longer time frame. However, the reproduction number can also be estimated from age-specific contact patterns. As we had launched the CoMix survey on the day of lockdown (24th March) we compared the contact patterns we observed in that first week to pre-COVID data the POLYMOD study to allow us to estimate what the reproduction number was likely to be **{JE/45 - INQ000092682}**. Reassuringly, the data suggested that the change in contact patterns would result in the reproduction number being significantly below 1 **{JE/9 - INQ000092698}**. The initial results were reported verbally to SAGE immediately (on the 26th March) **{JE/100 - INQ000119726 }**. Reports were then prepared and submitted to SPi-M-O (on the 26th March) and then published on our repository **{JE/101 - INQ000212229 }**. The peer-review paper was published in May 2020 **{JE/52 - INQ000092690}**.

8.25. Some observers, such as Dr Richard Horton, have indicated that SAGE appeared to be too slow to recommend action during the early weeks of the epidemic. I have some sympathy with this view and indeed was becoming increasingly anxious during this period as it was not obvious to me to what extent the Government and responder organisations (such as the NHS) were preparing for COVID-19. Ordinary SAGE attendees were not party to discussions at senior levels, and we were not tasked with,

or constituted to look at, operational matters. Many of us were relieved when Dominic Cummings first attended SAGE (on March 5th 2020 **{JE/102 - INQ000106152}**) as it indicated that the Government was taking this issue seriously. Also, SAGE did not offer explicit advice on the "best" course of action. It could not do this, as it was only looking at one narrow (scientific) point of view. Instead, it offered insights into the potential impact of different interventions and left the decision-making to those who were empowered to take these decisions. This attempt to remain neutral with respect to policies also meant that the minutes of the meetings were rather dry. Any urgency in the discussions was lost and differences in tone were neutralised. Finally, as I have previously said, it was unclear (at least to me) what the acceptable political limits were for different policies. Was lockdown a serious option before mid-March 2020? Would border closure have been acceptable in January 2020? I very much doubt it, but I don't know. Looking back, should we have been more proactive? Again, I don't know. We were clearer in advocating a policy course regarding the Autumn wave in September 2020. However, the Government did not follow this advice at the time which caused significant tension between SAGE advisors and the Government. It is certainly possible that stronger recommendations during the early phase of the pandemic might have led to faster action. But it is equally possible that such an approach might have been met with a similar response and endangered the system of scientific advice at a critical time.

8.26. Similarly, I felt that the messaging in January and February 2020 was very reassuring. I presumed that this was deliberate. I thought it best that these messages were handled by senior civil servants and politicians. Nevertheless, I did think that at times the messaging was potentially reassuring. One such example was in relation to the PHE risk assessment regarding Covid variants of concern, signed off by NERVTAG on 21st February 2020 **{JE/103 - INQ000119469}**. The PHE risk assessment had a very narrow, specific scope (it was intended to assess the current not future risk). However, I felt at the time that stating that the risk was "moderate" when all indications suggested that we were about to experience a very severe epidemic could be misconstrued and could lead to complacency amongst decision-makers.

8.27. SAGE did not discuss overall strategy – planning was an operational issue, and setting the overall direction and goals was for the Government to decide. Thus, when strategy documents were released – such as the one on 3rd March 2020 – then this was the

first time that the ordinary SAGE attendees saw them. The plan published on 3rd March 2020 (contain, delay, research, mitigate) **{JE/104 - INQ000212035}** implied that the goal was to mitigate the epidemic, rather than to suppress it – although one could argue that suppressing transmission is a form of mitigation. I felt that a decision had been made and our job - as mathematical modellers - was to attempt to sketch out the implications of different mitigation options. This is what we did during the first weeks of March 2020 (and indeed, beyond). As we accumulated more evidence on the possible effectiveness of different options it became clear to me that we had to take much more radical measures to mitigate the epidemic than was being discussed at the time (there was a lot of discussion about banning sporting events, but it was clear that this would do very little on its own). It was also not clear to me what preparations had been made in the health and care sectors (as well as elsewhere, such as schools) to cope with the coming wave and the possible mitigation measures. My replies to Sir Jeremy Farrar's emails on 12th March indicate my alarm and frustration at the time and the need to put significant additional measures in place very quickly to avoid a disaster **{JE/105 - INQ000212036; JE/106 - INQ000212037; JE/107 - INQ000212038}**.

8.28. A prevalent view on SAGE at the time was that a summer wave of COVID-19 would be preferable to a winter one, which would carry with it the additional risk of other seasonal viruses, putting further pressure on stretched health care resources. My view was that we should concentrate on the immediate and much more pressing problem of a major epidemic of COVID-19 and worry about other seasonal pressures if and when necessary. In particular, I thought that if we had enough measures in place to control COVID-19, then it would be very unlikely that we would have an influenza epidemic, given COVID had a reproduction number of about 2.5-3 and seasonal influenza typically has a reproduction number of about 1.5 or lower. More importantly, I also felt that attempting to "get the epidemic over with" by winter would necessarily cause a very large wave in the summer. I felt that we had to reduce and spread out the wave over a much longer period – at least a year or more – in order to stop the NHS from being overwhelmed by COVID-19 cases. This would also potentially buy enough time for vaccines or better treatments to become available.

8.29. It was also clear that imposing strict restrictions and then lifting them would result in a bounce-back in cases (as noted by SAGE on 13th March 2020 **{JE/84 - INQ000212212}**). That is, imposing a lockdown and then returning to normality would

delay the epidemic but not avoid it. So, the benefits of a lockdown on its own were not clear unless it lasted until another technical solution (e.g. vaccines) could be found. Given the common belief that widespread vaccination might be available in 12-18 months at best, this seemed hard to imagine being acceptable. On the 11th March 2020 we came up with a potential solution to this **{JE/91 - INQ000212040}**. The idea was to have a background of social distancing measures in place for a long time (we initially modelled these as being in place for 7 months but extended this in later analyses, see for example **{JE/108 - INQ000212041}**. These measures would slow transmission (and offer some protection to the elderly) but would not be sufficient on their own to suppress the epidemic (i.e. reduce R to less than 1). However, on top of that, periodic lockdowns could be put in place when pressure was starting to build on the health service. The threshold for intervention could be adjusted – the lower the threshold the less pressure on health services and the fewer the resulting deaths, but the longer time would be spent under severe restrictions. These lockdowns would temporarily suppress transmission and reduce the pressure on the health service. In this way the epidemic would be kept at a low level, and thus spread out over a long time. This could buy enough time for other technological solutions to become available (e.g. vaccines or therapeutics). However, even if these were not forthcoming, keeping the epidemic at a low level would greatly reduce the deaths from COVID-19 (and prevent the NHS from being overstretched). The reasons for this were both simple – keeping infections low results in few deaths – but also complex. The complicated factor is that epidemics overshoot. At the peak of an unmitigated epidemic, the amount of immunity in the population is such that chains of transmission become broken and the epidemic starts to decline (R is less than 1, that is, each case generates, on average, less than one other case). However, at this point there is typically large numbers of infectious individuals, so many more cases occur as the epidemic declines from a high peak. Indeed, there may be almost as many cases in the declining phase of the epidemic as in the increasing phase. So, keeping the numbers of cases low at all times reduces the total number of infections and deaths by avoiding this overshoot. These initial ideas were first tabled at SPI-M-O on March 11th and were tabled again at SAGE on March 13th **{JE/109 - INQ000228589}.** Professor Ferguson also looked at similar strategies in his influential paper of the 16th March **{JE/110 - INQ000212042}**. In the end, the overall approach that we adopted over the course of the epidemic was close to this strategy – a background of non-pharmaceutical measures that slowed but did not

suppress transmission on top of which we instigated periodic lockdowns to reduce incidence and protect the NHS. It stretched the epidemic out until we had widespread vaccination in place, after which we could begin to relax measures. The main difference was that we adopted this periodic lockdown approach in an ad-hoc and responsive way, instead of a deliberate and planned one. Also note that this approach can be regarded as either a mitigation strategy or something akin to complete suppression, depending on the threshold for lockdown that is adopted. The acknowledgement that cases will rise again after easing of a lockdown does not necessarily imply that we are following a mitigation strategy, and in any case, it is possible to adopt a low incidence mitigation strategy, as many countries did.

9. The timing of the first national lockdown

9.1. The decision to put the country into "lockdown" in March 2020 was one of the biggest peace-time decisions any Government has had to make. Although the epidemiological and modelling evidence was emerging by that time that lockdown was likely to be effective, the evidence for the ineffectiveness of alternative (less stringent) measures was not as clear. The knock-on (indirect) effects of lockdown on physical and mental health were unknown, as were the wider economic and social costs. They were, however, likely to be very substantial. There is therefore a danger, when questioning the timing of the lockdown in retrospect, to assume that this decision was obviously the correct one. I happen to think that going into lockdown was the correct decision and I wish that we had done it earlier. However, at the time the Government did not have the benefit of experience of these measures or of hindsight.

9.2. I can only speculate as to the factors that might have led to the Government taking the decision when it did, rather than earlier (or indeed later). Many different factors must have played a role, including economic, operational, and political ones. I have no knowledge or expertise in these areas and will not comment on them further. I will confine my comments to the scientific and epidemiological issues that may have played a role.

9.3. The Institute of Government concluded that the Government's wish to avoid lockdown was a factor. This is certainly possible. It took time to explore alternatives. The Imperial and LSHTM models could start to be used for scenario analyses in late February and

early March (other groups associated with SPI-M-O took a little longer). Over the first few weeks of March various scenario analyses were undertaken to assess the impact of alternative policies. However, this takes time. Each time a question was posed, it would take a few days to get an answer and then to report this back to SPI-M and then to SAGE. It is inevitable that there are some delays inherent in this process. I would also query some of the prioritisation at this time. For instance, we were repeatedly asked about the impact of banning sporting events. Models were probably not necessary to help answer this question. Indeed, given the models are population-based and the impact of banning sporting events at the population level was likely to be small, it is conceivable that the modelled answer did not help the decision-making.

9.4. A second issue is the data that we were working with at the time. As explained in an earlier section, the surveillance system was poor, with the data being delayed and hard to interpret – so much so, that estimating the growth rate of the epidemic was difficult (estimates of the doubling time varied from about 3 days to about 5-7 days, depending on what method and data sources were used) and getting an accurate assessment of the overall size of the epidemic was also difficult. The delays and under-reporting (partly due to a lack of testing) might well have led decision-makers to conclude that they had more time to act than was the case. Indeed, I had brought this up at the SAGE meeting of 13th March **{JE/84 - INQ000212212}** and wrote to SAGE's Co-Chairs on this issue on Sunday 15th March 2020 **{JE/111 - INQ000212043; JE/112 - INQ000212044; JE/113 - INQ000212045; JE/114 - INQ000212046; JE/115 - INQ000212047; JE/116 - INQ000212048; JE/117 - INQ000212049; JE/118 - INQ000212050; JE/119 - INQ000212051; JE/120 - INQ000212052; JE/121 - INQ000212053; JE/122 - INQ000212054; JE/123 - INQ000212055; JE/124 - INQ000212056; JE/125 - INQ000212032}**. Surveillance started to improve after the CHESS **('COVID-19 Hospitalisation in England Surveillance System')** system was launched in hospitals on around 14-15th March, but it took a little while for the new data-stream to stabilise. Nevertheless, by the third week of March it was becoming clearer that ICU capacity would become under severe strain in some areas (notably London) by the end of March and that because of inherent delays between infection and hospitalisation, there was very little time to act to stop a crisis.

9.5. Third, the optimal timing of interventions depends on the goal of the intervention and the operational constraints acting on its implementation. It is usually better to act as

rapidly as possible due to exponential epidemic growth. Delays in reporting and under-reporting also imply that it is usually better to err on the side of wider and more stringent actions as the epidemic may be significantly larger and more widespread than the data imply. However, there can be exceptions to this these general rules. As the models had shown by early March 2020, very stringent population-wide non-pharmaceutical interventions could potentially reverse epidemic growth. However, such interventions would need to be in place until some other measure could take over the role of preventing transmission, the most likely measure able to do this being mass immunisation with an effective vaccine. If the non-pharmaceutical interventions can only be in place for a few months (for whatever logistical, political, economic, or social reason) then simply releasing them will lead to a surge of cases again. That is, stringent early interventions followed by return to normality will delay the epidemic rather than avert it. If measures can only be put in place for a short period, then more cases can be averted by putting those measures in place later to reduce the peak height. Scenario analyses conducted in early March concentrated on short interventions (a few months). There was no indication that longer interventions might be acceptable, or indeed that severe restrictions would be politically palatable at the time (see earlier comments about the lack of feedback about what the limits of acceptability were).

9.6. To what extent different interventions (and their duration) might be acceptable to the public was also largely unknown. Western countries had not implemented widespread emergency non-pharmaceutical interventions for more than 100 years. Apart from some polling, there were no observational data on their acceptability and whether this might be affected by the duration of the intervention. Thus, the public acceptability of long-term stringent interventions was – as far as I was aware – not known at the time.

9.7. It was therefore not at all clear whether lockdowns were feasible and acceptable – politically, economically, or socially. To my mind at least, the Italian lockdowns of 8-9th March changed this with the possibility of a lockdown then being discussed in the press (up to this point there had been remarkably little public debate about the pandemic). SPI-M-O started looking at the consequences of lockdown on 10th March. It was not asked to do so.

9.8. SPI-M-O's work on lockdowns was stimulated by a short note by Professor Steven Riley written on the 10th of March that resulted in an attempt to write a consensus

statement that covered the major options (suppression until a vaccine or mitigation). I offered a third option (as discussed earlier) which involved flattening the epidemic far more and for much longer than had been considered before, by using long-term social distance measures interspersed with intermittent lockdowns. This was modelled by our team and that of Professor Mark Woolhouse on the 11th of March, discussed by SPI-M-O and brought to the attention of the CSA **{JE/126 - INQ000212057; JE/127 - INQ000212058; JE/128 - INQ000212059; JE/129 - INQ000212060; JE/130 - INQ000212062; JE/131 - INQ000212063; JE/132 - INQ000212065; JE/133 - INQ000212067; JE/134 - INQ000212068; JE/135 - INQ000212070; JE/136 - INQ000212071; JE/137 - INQ000212075}**. This was the first time that we had seriously considered lockdown strategies, including ways to exit lockdowns. Professor Ferguson built on these early results and analysed similar strategies in his paper on 16th March **{JE/110 - INQ000212042}**.

9.9. The decision to enter lockdown and the timing of that decision was a political one. The Government had to weigh the multitude of potential benefits and costs of this option against alternatives. Many factors needed to be taken into account, not just scientific ones. From the epidemiological point of view, the poor quality of the surveillance data (with only tens of cases being reported on a daily basis) – **{JE/111 - INQ000212043; JE/112 - INQ000212044; JE/113 - INQ000212045; JE/114 - INQ000212046; JE/115 - INQ000212047; JE/116 - INQ000212048; JE/117 - INQ000212049; JE/118 - INQ000212050; JE/119 - INQ000212051; JE/120 - INQ000212052; JE/121 - INQ000212053; JE/122 - INQ000212054; JE/123 - INQ000212055; JE/124 - INQ000212056; JE/125 - INQ000212032}** and the relative lack of work on lockdowns by the modelling groups meant that the scientific evidence-base only began to emerge during the second week of March to help guide decisions on lockdowns or their alternatives. The Government could have decided without detailed scientific evidence (other countries clearly did) but if they were waiting for the evidence to emerge, then this did not really start to accrue until the SAGE meeting of 13th March **{JE/84 - INQ000212212}**. An announcement from the Government was expected on 16th March. I have always felt that this is the first feasible date that a decision to go into lockdown could have been backed-up by UK-specific scientific analysis. Lockdown was announced one week later on 23rd March 2020. As the epidemic was increasing rapidly at the time, this delay almost certainly led to thousands of additional deaths. It

is hard to calculate how many, as behaviour appeared to be changing even without lockdown, but subsequent analysis suggested that the timing of the lockdown, relative to how large the epidemic was at that point, was a major determinant of the death toll in the first wave **{JE/138 - INQ000212076; JE/139 - INQ000212077; JE/140 - INQ000212078}**.

9.10. It is certainly possible that had SAGE been earlier, clearer and more urgent in its advice then lockdown could have been introduced earlier. To offer the evidence base for earlier and clearer advice would have required earlier analyses and better data. Surveillance was transformed after the lockdown with an enormous injection of resources. If this revolution in testing and surveillance was being considered earlier (i.e. by early March), then I am not aware of it. Likewise, we did not model lockdowns in a detailed way (including thinking how to exit from them) until mid-March. This could have been brought forward by about a week or so (model results were very uncertain before then). However, SPI-M-O was not tasked with looking at lockdown strategies at this time and was fully employed looking at other – less stringent - measures. Indeed, SPI-M-O members started to address this issue on the 10th-11th March without being asked to do so. Given the lack of a request, it seems unlikely that the Government were really considering these options until later.

10. April 2020 onwards

April to May 2020

10.1. The first national lockdown was highly effective at reducing transmission in the community. The measures taken on the 23rd of March 2020 reduced the reproduction number to significantly below one (though it took some time for this to be confirmed by the epidemiological data) and cases declined. It is possible that individuals will have changed their behaviour to reduce their risk – as was the case in Sweden – but it is unlikely that this would have been as rapid or as effective as lockdown was (comparing Swedish case data with that of neighbouring countries is strongly suggestive of this).

10.2. As cases peaked in the community, the epidemic in other settings took hold – primarily hospitals and care homes, though other enclosed settings such as hostels and prisons were also at risk. The hospital and care-home epidemics led to very considerable illness and loss of life. Roughly half of all COVID deaths in the first wave were attributed

to care home residents (ONS reported 19,783 COVID-related deaths in the first wave in England and Wales, but the true number was almost certainly higher than this **{JE/141 - INQ000212079}**). It had been clear from the earliest data available on the clinical impact of COVID-19 that the elderly and those with certain pre-existing conditions were at highest risk (see for example: **{JE/142 - INQ000212080}**). That hospitals and care homes were potential high-risk environments was not a surprise. It is clear that not enough was done in February and March 2020 to reduce this risk.

10.3. The care home epidemic was well-recognised at the time, and rightly so. The epidemic in hospitals was much less discussed. It was, however, very significant. Regular reports from the CoCIN study were provided to SAGE during this time. These used a simple measure of probable hospital-acquired infection, based on the onset date. If the patient first showed signs of COVID-19 illness at least 5-7 days after they were admitted, then they probably acquired the infection in hospital as the mean incubation period was about 5 days (other thresholds were also used as sensitivity analyses). This simple metric suggested that almost 10% of hospitalised patients with COVID probably acquired it in hospital by the end of March 2020 **{JE/143 - INQ000212081}**. This measure was simple to calculate and track, but it was also an underestimate of the true burden of hospital acquired infections as the average length of stay in acute NHS hospitals was only about two days. Hence, about 2/3rds of hospital-acquired cases would be discharged before they displayed symptoms. Accounting for this suggests that there were over 26,000 hospital acquired symptomatic cases in the first wave of COVID-19 in England **{JE/144 - INQ000212082}**. To put this in context, there were roughly 110,000 hospitalised cases of COVID-19 in England during the first wave (until the end of July 2020).

10.4. SAGE was not tasked with operational matters; these were dealt with by the appropriate organisation (e.g. the NHS). SAGE offered scientific evidence (e.g. on the scale of the problems, see above) and on the possible effect of mitigation measures, such as testing. SAGE members (particularly the CMO and CSA) were also influential in setting up large-scale studies in hospitals and care homes around this time **{JE/145 - INQ000212083}**. Nevertheless, tracking these epidemics in our most vulnerable groups was both alarming and, at times, frustrating for many of us on SAGE **{JE/146 - INQ000212084}**.

10.5. Other large-scale studies were set up at this time to track the epidemic, understand transmission and the effectiveness of different measures as well as track the evolution of the virus. These studies, together with the enormous expansion of testing that occurred in April 2020, transformed our surveillance of the epidemic.

10.6. This period also saw the results of randomised controlled trials of different therapeutics start to become available, including the RECOVERY trial led by Professor Peter Horby. These studies led to improvements in treatment for COVID-19 cases. Indeed, the fraction of hospitalised cases that died fell from around 35% in early April to around 20% in early summer, although it is likely that much of this decline was due to easing of hospital pressures, as the hospital fatality ratio started increasing again in the autumn wave.

10.7. Traditional contact tracing is very resource intensive. Cases need to be kept low for contact tracing to be effective. If resources are stretched so that contacts are not rapidly traced and asked to quarantine, then a feedback loop can build up leading to more cases and more pressure on the system. Digital (app-based) contact tracing does not have such problems, but still requires a system for rapid diagnostic testing to be in place to be effective. With the decline of cases, it became feasible that contact tracing could be used to replace some of the national level restrictions. That is, population-wide measures would be replaced by targeted measures aimed at those who were most likely to be infectious to others. A revamped and enormously expanded test, trace and isolate system was to be launched at the end of May. There are many different ways of delivering contact tracing. SPI-M-O and NERVTAG were tasked with evaluating different options, aided by modelling work from CMMID **{JE/22 - INQ000092657; JE/22A – INQ000255418}**. This work updated and extended earlier work on contact tracing **{JE/11 - INQ000092645; JE11A – INQ000255406}** to use UK-specific data on contact patterns, updated epidemiological parameters and realistic estimates of operational delays (guided by NERVTAG). It also looked at combinations of test, trace and isolate policies alongside other NPIs, and found that contact tracing could hold the reproduction number below one, if a high level (>80%) of contacts outside the household were quickly traced and moderate NPIs remained in place. This work helped SAGE set recommended performance targets for the Test, Trace and Isolate system **{JE/147 - INQ000120511}**. The NERVTAG/SPI-M-O committee also looked at different modes of operation. It recommended, for instance, that quarantine

of contacts should be initiated on a symptomatic case and should not be delayed until laboratory confirmation. If the suspect index case was found to be negative, then the contacts should be released **{JE/148 - INQ000120452}**. This was not adopted when the new TTI system was implemented.

10.8. One other concern for me at the time of the launch of the test, trace and isolate system was whether it would be able to cope with the incidence of both COVID and other infections. As the ONS Coronavirus Infection Survey had been launched by then, we had a good estimate of the prevalence: about 1 in 600 people were testing positive, which would equate to roughly 100,000 infections nationwide. That seemed like a lot of infections to test and contacts to trace. In addition, however, as the clinical definition was broad, other infections would also trigger an investigation. That is, the system had to be sufficiently resourced to cope with cases of COVID-19 (which were still relatively common) and other infections that could be mistaken for COVID-19 **{JE/149 - INQ000212088; JE/150 - INQ000212090}**.

10.9. Finally, we did not know how people would react after restrictions were lifted. This had never been done before so there was little or no evidence to draw on. Most people assumed that individuals would return to normality (i.e. pre-pandemic levels of mixing) as soon as they were allowed to. This could lead to a surge of infections which would put the Test and Trace system under substantial pressure.

10.10. Together with the fact that NHS Test and Trace was an entirely new entity, I was nervous that too much was being expected of it and it would not be able to successfully keep incidence low. I also thought that it was best that incidence should remain as low as practicable, as this was the best way to minimise deaths, protect vulnerable and elderly individuals and allow the NHS to function as normally as possible. For all these reasons, I thought that we were easing restrictions too early in late May 2020.

Planning for the future

10.11. On April 10th, 2020, a small group discussion was organised by SAGE on the "Science of Exit (from lockdown)". This was unusual in that it was a small group of SAGE participants and that the discussion concerned strategy. From memory, those present included the CMO and CSA, Dame Angela McLean, Professor Graham Medley Professor Neil Ferguson, Professor James Rubin, Professor Brooke Rogers and

myself. I cannot recall all of the attendees. The discussion crystallised around the merits or otherwise of trying to keep incidence as low as possible or allowing infection levels to be relatively high but still within the bounds of what the NHS could cope with. Following the Low Incidence scenario would mean remaining under restrictions until an effective vaccination programme had been rolled out, whereas under the Higher Incidence scenario infections would be allowed to accumulate. Thus, natural immunity would build in the population, eventually allowing restrictions to be lifted even if a vaccine had not been developed. Clearly this latter option resulted in far more infections and therefore deaths and would have an adverse effect on the health service as hospitals would remain busy with COVID-19 cases over an extended period of time. In sketching out these two broad scenarios, two things became apparent. First, to hold the incidence at a low level or to hold the incidence at a high level required approximately the same interventions (I.e. in both scenarios the reproduction number was 1 (each case generated one more case) and the same package of non-pharmaceutical interventions would be necessary). That is the steady state cost of interventions in both scenarios was approximately the same. The main difference was that restrictions would be eased a few weeks before in the Higher Incidence scenario. Second, even under the "Higher incidence" scenario levels of infection would have to remain relatively low to avoid the NHS from being overwhelmed. This meant that to achieve "herd immunity" through this Higher Incidence scenario would likely take 1-2 years. This is roughly the same time that a vaccine might be expected to be available and rolled out, so both scenarios might well result in restrictions for a similar amount of time. So, the restrictions necessary and thus the social and economic costs of both scenarios were likely to be similar and they were both likely to last a similar length of time (1-2 years). However, the Higher Incidence scenario would cause many tens of thousands of more deaths and result in the NHS being under sustained pressure over a long time. These scenarios were sketches and the choice was ultimately a political one. However, it was very clear to me which was the better course of action **{JE/151 - INQ000212091; JE/152 - INQ000212092; JE/153 - INQ000212094; JE/154 - INQ000212095; JE/155 - INQ000212097; JE/156 - INQ000212099; JE/157 - INQ000212100}**.

10.12. I do not know to what extent these discussions and this document were used within Government. Subsequent events suggest that the message was not heeded.

Easing of restrictions in the summer of 2020

10.13. Restrictions were eased over the late spring and early summer of 2020. The order of the lifting of restrictions and the timing of the lifting was (as ever) a political decision. There was very little evidence at the time regarding the risks associated with different settings (hairdressers, gyms etc) and the effectiveness of the different measures that were in place: they had been implemented together in March 2020 and so it was not possible to tease out the separate epidemiological effect of the different components of the interventions. Lifting them in a more phased way, with a period for evaluation between each stage, could have helped determine how effective different packages of interventions might be.

10.14. The final package of restrictions was lifted on the 4th July 2020. Cases started to increase immediately afterwards, though this was not apparent initially as case numbers were low and there were significant day effects (i.e. variation within a week because of differences in testing and reporting on different days of the week). By the end of July, it was clear to me that there was a consistent trend upwards in the cases and I brought this to the attention to others in an addendum to an email to the CSA and others on 27th July **{JE/158 - INQ000228590}**.

10.15. We were also closely monitoring contact patterns through the CoMix survey. They were increasing, though very gradually. Individuals appeared to be acting far more cautiously than had been anticipated. Mean contact rates remained far below pre-pandemic levels.

10.16. The messaging from Government had also started to get more confusing, with ministers suggesting we should start to return to work (the messaging around this time was famously mocked by the comedian Matt Lucas). Then in August of 2020 "Eat Out to Help Out" was launched. I thought that this was misguided in the extreme and would help to kick-start a second wave. It was not the only measure, as the pressure to return to work was starting to be applied, along with the general opening up of society. I felt that it was one thing to permit opening of various sectors of the economy, but quite another to actively encourage risky behaviour. "Eat Out to Help Out" did exactly this, with public funds being offered to individuals to take an epidemiological risk (the co-financing only applied if customers ate in – it did not apply to take-aways). This was

actively throwing away the strategic advantage that we had gained over the epidemic. We had driven cases to low levels, hospitals had started to return to normal, the epidemic in care homes was much reduced and yet the Government was now taking active steps to reverse the gains that the country had made at such huge economic and social cost.

10.17. In addition to this, we had opened up travel in the summer of 2020, with quarantine-free travel permitted to many European holiday destinations. This led to the importation of a new, slightly more infectious variant (first identified in Spain) that gradually became dominant in the UK over the autumn. Border restrictions were useless during the spring of 2020 when we were experiencing tens of thousands of endemic infections per day. Opening up the borders when we had driven infections to low levels made little epidemiological sense.

10.18. The flip-flopping of policy objectives, from suppressing the virus to encouraging (even paying for) epidemiologically risky behaviour showed a lack of strategic thinking and planning. Did the Government have a long-term objective and a plan as to how to achieve it? It seems not. It is hard to believe that this inconsistent approach minimised the economic and health cost of the pandemic.

10.19. With schools and universities about to open, the epidemic increasing, and pressure being applied to return to work, the prospects for the autumn appeared rather bleak.

Autumn 2020

10.20. On 10th September I was asked by SAGE to chair a working group to review the possible epidemiological impact and social and health harms of reintroduction of different NPIs. This was brought back to SAGE the following week and after further comments from the committee was reconsidered and approved by SAGE on 21st September **{JE/159 - INQ000212102}**. The paper and associated table summarised the evidence available at the time but has often been characterised simply as SAGE calling for a circuit-breaker (short) lockdown **{JE/160 - INQ000212103; JE/161 - INQ000212104}**. This was not the case. The paper, endorsed by the committee, suggested a circuit-breaker to reduce incidence, but also the reintroduction of other measures over the longer term to keep incidence low. I have copied (below) the first two bullet points of the SAGE minutes from September 21st **{JE/156 - INQ000212097}**

10.20.1. COVID-19 incidence is increasing across the country in all age groups. The effect of opening of schools, colleges and universities has only just begun to affect this increase. Even so, the latest data suggest that the doubling time for new infections could currently be as short as 7 days nationally. COVID-19 related hospitalisations and intensive care bed usage have started to rise. SPI-M has modelled the potential increases.

10.20.2. A package of interventions will need to be adopted to reverse this exponential rise in cases. Single interventions by themselves are unlikely to be able to bring R below 1 (high confidence). The shortlist of non-pharmaceutical interventions (NPIs) that should be considered for immediate introduction includes:

10.20.2.1. a circuit-breaker (short period of lockdown) to return incidence to low levels;

10.20.2.2. advice to work from home for all those that can;

10.20.2.3. banning all contact within the home with members of other households (except members of a support bubble);

10.20.2.4. closure of all bars, restaurants, cafes, indoor gyms, and personal services (for example hairdressers); and

10.20.2.5. all university and college teaching to be online unless face-to-face teaching is absolutely essential.

10.21. I was asked to attend a meeting with the Prime Minister on Sunday 20th September. The Chancellor (Mr Rishi Sunak MP) was also present, as were a few other officials (who did not speak) and a few academics and public health experts, namely Professor Sunetra Gupta, Professor Carl Heneghan, Dr Anders Tegnell and Dame Angela McLean. The idea appeared to be to listen to a range of views. We were asked to produce a one-page assessment of the epidemiological situation at the time **{JE/162 - INQ000212105; JE/163 - INQ000212107; JE/164 - INQ000212108; JE/165 - INQ000212109}**. I tried to point out that the epidemic was increasing exponentially and that harsh measures would have to be introduced soon to stop the NHS from being overwhelmed. That is, the decision was not to lock-down or not, but to lock-down now, or to be forced into locking-down later. The former strategy would minimise deaths and disruption, the latter would lead to unnecessary deaths, pressure on the health system and would necessitate longer and/or more stringent measures to be introduced later

to bring the epidemic back under control. I pointed out that the UK had reacted late in March and had paid a high price for this both epidemiologically and economically and that it would be wise to learn from this recent history. My arguments were clearly not persuasive enough.

10.22. Meanwhile the epidemic continued to increase, and the testing programme started to come under significant pressure, with shortages and queues at many testing centres. In response, the "Rule of 6" had been introduced on September 14th **{JE/166 - INQ000212110}**. It had not been discussed by SAGE. Instead of following SAGE's advice to introduce a comprehensive package of interventions, the Government introduced a 10pm curfew for bars and restaurants at the end of September. Then, on October 14th, tiered restrictions came into place. Cases and hospitalisations continued to increase over this period.

10.23. The tiered system was particularly badly thought-out. Restrictions in Tiers 1 and 2 were light and very unlikely to reverse epidemic growth. The restrictions under Tier 3 were more stringent and could potentially halt epidemic growth but were unlikely to be sufficient to reverse it (based on our assessment of the impact of interventions at the time – (see table attached to September 21st paper **{JE/161 - INQ000212104}**). Local authorities would be placed into higher restriction levels if the incidence rose. Given the weak restrictions in the lower tiers, local authorities would inevitably increase in incidence until they were in Tier 3 and then the incidence would be held roughly level. That is, instead of preventing the incidence from rising, it would be allowed to rise to the point that cases were high, hospitals were under strain and care homes were under increased risk. The epidemic would then be held at this high level. The system would eventually lead to all local authorities having high incidence. It was epidemiological levelling-up.

10.24. The epidemic continued to grow, and as hospitals came under increasing pressure the Government were forced (by the virus) to implement national-level restrictions. On the 31st of October the Government announced the second national lockdown, which was to start on the 5th of November and to last for 4 weeks. Schools were to remain open. Unlike in Northern Ireland and Wales, the measures were not timed to coincide with the school half term and so an opportunity to use this to an advantage was lost. Once again, policy appeared to be reactive, rather than planned.

10.25. Hospital infections had increased again over the autumn, as did outbreaks in care homes. Testing of staff and patients/residents was in place and there was much better access to personal protective equipment. More comprehensive testing (including of carers) and better isolation procedures could, perhaps, have helped protect the vulnerable. However, we could not (and perhaps should not) effectively isolate the vulnerable from society. The best way to protect them was to keep the incidence low.

10.26. The second lockdown was not as severe as the one in March (e.g. schools remained open) and cases declined slowly (R was estimated to be between 0.9 and 1 over this period **{JE/167 - INQ000212111}**. Thus, when the lockdown was eased at the beginning of December, cases remained high, and hospitals remained under pressure (there were 16,000 COVID patients in hospital at the beginning of December 2020). Had measures been put in place earlier in the autumn, we could have returned to the low incidence seen in the summer. By implementing the second lockdown so late, it was insufficient to return us to a low incidence scenario. It reduced the immediate pressure on the health service somewhat, but this benefit would not have lasted for long. Even without the Alpha variant (which no-one was aware of at the time) further measures would have been needed in the near future as pressure started to build once again.

10.27. There was some good news. On 9th November 2020 Pfizer and BioNTech released interim results of their pivotal vaccine trial, suggesting that the BNT162b2 mRNA vaccine was more than 90% effective at preventing COVID-19. The vaccine was granted temporary authorisation by the MHRA on 2nd December 2020 and the first person to be vaccinated outside a clinical trial occurred on the 8th December. The potential for vaccines to be available by the end of the year was known in the autumn of 2020 **{JE/168 - INQ000212112}**. That is, in the autumn of 2020 there was a prospect that the end of the epidemic might start to come into sight (the emergence of new highly transmissible variants later damaged this optimism).

10.28. At the end of the first wave (July 1st 2020) there had been 40,780 deaths within 28 days of a positive test for COVID-19 (other measures of death gave similar numbers). By 11th December there had been 65,376 (the surge of deaths due to the Alpha strain occurred after this date) **{JE/169 - INQ000212113}**. That is, roughly 25,000 people died in the Autumn wave before the Alpha variant started to take its toll, despite better treatments and better protection of our hospitals and care homes. The advantages of

early interventions were once again lost and so when the lockdown was eventually implemented it was then insufficient to return the country to a low incidence scenario. In the first wave, our surveillance was poor, and we were faced with a new disease that we had no experience of controlling. Decisions were taken too late, resulting in a major wave of infections. A lack of adequate measures in care homes and hospitals allowed the epidemic to spread in our most vulnerable groups, significantly adding to the death toll. By the time the second wave hit, our surveillance system was amongst the best in the world, with high levels of PCR testing and genomic confirmation; large-scale random testing studies (the ONS Coronavirus Infection Survey and the REACT study) tracked the underlying epidemic in the community and other major studies were underway in higher-risk settings such as care homes, hospitals, schools and households. The failure to act quickly to control this wave was not because of a lack of situational awareness or knowledge of how to control it. We let this second wave happen. The first national lockdown brought about low levels of infections, hospitalisations and deaths. It achieved this at enormous economic and social costs. Instead of capitalising on this situation we threw that strategic advantage away for a short-term economic boost – despite the nearing prospects of vaccines becoming available and the unambiguous and timely advice of the Government's own scientific expert group.

The Alpha Wave and Third National Lockdown

10.29. On August 1st 2020 there were 879 patients in hospital in England with COVID-19. By December 2nd there were 13,212 **{JE/169 - INQ000212113}**. Many public health experts had been warning about the winter for many months, as the NHS typically comes under pressure during this period. The failure to act promptly in the autumn meant that we were very poorly prepared. Although a wave of influenza and other infections did not materialise, a new more transmissible and pathogenic strain of SARS-CoV-2 did. I said at the time that it was the worst moment of the epidemic. This was for many reasons: the hospitals were already under pressure with COVID-19 cases; health and care workers had been working under enormous strain for months; the new variant was increasing rapidly despite the revamped (and strengthened) Tier system; and the Government had just demonstrated its reluctance to act quickly to take necessary public health measures.

10.30. In fact, the Government did act relatively quickly. It took a few weeks in early to mid-December to establish that the pockets of high incidence were due to the spread of a new variant. Estimation of the growth rate of the new variant and the extent that this was due to immune evasion or increased transmissibility was critical, as was the inherent pathogenicity (i.e. whether the infection fatality ratio was different). Early laboratory studies suggested that immunity targeted to previous strains was likely to be effective. However, the new variant had first been identified spreading rapidly in areas that were under Tier 3 restrictions (the highest at the time). This suggested that either those restrictions were not being followed in those areas or the virus was inherently more transmissible. The CoMix data suggested that there was no significant difference in contact patterns in affected areas compared with other parts of the country, implying that the new variant was more transmissible, which proved to be the case.

10.31. Careful analysis using multiple statistical and modelling methods suggested that the new variant was roughly 50-100% more infectious than the previous strains **{JE/35 - INQ000092671; JE/35A – INQ000255437}.** At this point (mid to late December 2020) there was no evidence that the new variant was more pathogenic than previous strains (although that assessment would later be revised **{JE/57 - INQ000212186}**. Modelling suggested that a very large surge of infections, hospitalisations and deaths would result, despite the current measures and the start of the immunisation programme. Further, very stringent restrictions would be necessary along with an acceleration of the vaccine roll-out. These results were first presented at an extraordinary joint meeting of NERVTAG and SPI-M-O on 21st December 2020 **{JE/170 - INQ000212114}**. The Government had already introduced a fourth Tier on 19th December. A third national lockdown came into effect on 6th January.

10.32. Although the Government did act relatively quickly in this instance, there was some confusion and delay. There appeared to be some reluctance to take widespread measures before the Christmas period and schools initially opened in January 2021, only to close the following day.

10.33. The measures were effective. Hospitalisations peaked in mid-January, though by this time roughly 40% of all the beds available within the NHS were being taken up by COVID patients. Deaths within 28-days of a positive SARS-CoV-2 test increased from 65,376 on 11th December 2020 to 127,651 on 1st April 2021 **{JE/169 - INQ000212113}**.

That is roughly 60,000 people died in this wave, despite the start of the vaccination programme and the imposition of severe restrictions.

Roadmap out of lockdown and the Delta wave

10.34. The Roadmap out of Lockdown was a plan to lift the third national lockdown in a number of steps during the Spring and early Summer of 2021. The vaccination programme was ongoing during this period, so the restrictions were planned to be eased as a greater fraction of the population had been vaccinated (the highest risk groups were vaccinated first). There was not a commitment to keep the reproduction number below one, so relatively high levels of infection were expected to occur as measures eased. On the advice of SAGE there was a 5-week gap between most of the steps. This was intended to give enough time for an evaluation of the impact of the previous easement measures before the next step was taken. In practice, even with this gap, there was insufficient time to accurately estimate the impact of the previous step before the next had to be announced (due to the delayed effect of measures on outcomes such as reported cases and hospitalisations and the need to provide the assessment of the effectiveness of the last easement to the government one week before the next planned date). Nevertheless, the final step was delayed by a month (partly because of the emergence of the even more transmissible Delta virus) so that it would coincide with the end of the school summer term when transmission might be expected to decrease, given that school children were unvaccinated and a potential source of infection.

10.35. At each step SPI-M-O assessed the possible impact of the previous easement and medium-term projections of what might be expected following the next step. In alignment with usual SPI-M-O procedures there were multiple groups working independently on these model projections, in this case three groups: Imperial College, London; Warwick University; and LSHTM. The model results were discussed and compared and summary statements by the committee were prepared at each stage and fed back to SAGE. LSHTM's assessments are given here: **{JE/62 - INQ000212190; JE/171 - INQ000212115; JE/172 - INQ000212116; JE/173 - INQ000212117}**. There were key uncertainties, particularly around the impact of vaccination on severe disease and infection (efficacy estimates had to be re-estimated

as the Delta virus became dominant), the duration of immunity and the behavioural response to an easement of restrictions. Nevertheless, the Roadmap out of lockdown represented a sensible approach to gradually easing restrictions, informed by the best available evidence at the time. It demonstrated how much policy decisions had matured and improved by that time. More could have been done to establish the science-base underpinning NPIs and testing strategies, which could have greatly improved our strategy for living with the virus in the future. For instance, I advocated, undertaking a series of randomised controlled trials of testing strategies and NPIs with linked economic analyses. Although more could have been done had there been sufficient will, the easing of these restrictions was undertaken far more successfully than it had been previously and data were collected while doing so (the Events Research Programme was set up to evaluate the opening of cultural and sporting events, for instance) **{JE/174 - INQ000212118}**. The period was characterised by policy-makers effectively harnessing the scientific resources at their disposal to help improve decisions.

10.36. After the final national restrictions were eased in July 2021, the epidemic settled at a relatively high level over the autumn of 2021 (compared to our neighbours, for instance). A number of factors contributed to this, including relatively low levels of vaccine coverage in younger individuals and lower levels of restrictions compared to our neighbours, particularly in unvaccinated individuals (according to the Oxford stringency index: **{JE/175 - INQ000212119}**. From 25th May 2021 to 1st December 2021 deaths within 28 days of a positive test increased from 128,520 to 146,541. That is, there were about 20,000 additional deaths during the Delta wave. Our surveillance system remained excellent. The government chose not to impose further restrictions (such as mask wearing, working from home or vaccine passporting) over this period, though many of our neighbouring countries did.

The first wave of the Omicron variant

10.37. A new variant (later named Omicron) was first identified in South Africa in November 2021. It quickly spread to the UK. The growth rate of this new variant was alarming – the epidemic was doubling every 2-3 days, despite high levels of immunity in the country, as it had in South Africa as well. This stimulated an enormous national and

international effort to try to characterise this new virus in terms of its transmissibility and ability to evade the immune response. CMMID provided an early warning of the potential for this strain to spread within the UK to SPI-M-O and the CMO and CSA in early December and released a report on this on 11th December 2021 **{JE/176 - INQ000212120}**. There was, at the time, enormous uncertainty surrounding key epidemiological parameters, so the report focussed on 4 scenarios. In each scenario, we expected infections to reach levels not previously seen before (and this is indeed what happened). In addition, all of the scenarios (even the worst case) expected deaths to peak at a lower level than was observed in the Alpha wave (when vaccines were only just starting to roll out). Three of the four scenarios had hospitalisations peaking close to, or below, the Alpha peak, and the worst case scenario had hospitalisations peaking significantly above the Alpha wave **{JE/177 - INQ000212121}**. Our later assessment (23rd December) was more pessimistic, as the growth rate was estimated to be even higher by that point **{JE/177 - INQ000212121}**. There was no way to tell these scenarios apart in mid-December 2021. The Government decided to re-introduce minor restrictions ("Plan B") and extend the vaccine booster programme to younger ages.

10.38. The timing of the Omicron wave was fortuitous – we were completing the booster programme for the higher risk groups (the over 50s were being vaccinated at the time) and we had sufficient vaccine stocks and the necessary system in place to rapidly expand this to other age groups. In addition, lateral flow tests were readily available and were being widely used to help reduce risk of transmission to others **{JE/178 - INQ000212122}**.

10.39. Had the virus been more pathogenic than it proved to be there was a significant risk that the NHS would have been placed under major strain again. The speed of spread of the Omicron wave (and the delays to hospitalisation and death) meant that actions had to be taken early to avoid a wave of severe cases. Waiting for confirmation about the severity profile of Omicron-related disease would not have given time to act. It took about one month from first identification of the virus to the first quantification of its inherent pathogenicity. There were anecdotes from South Africa that it was less pathogenic than Delta before this time, but this was impossible to verify and South Africa had very high levels of immunity in the population. That infection would appear milder in previously infected or vaccinated individuals was expected and taken into

account in the models. The question was: is Omicron inherently less pathogenic, and so would the number of severe cases be lower than was modelled? Analyses to show this were not available until just before Christmas 2021 **{JE/179 - INQ000212123}**. Had it then been necessary to impose major NPIs at this point, then it would have been too late to avoid a significant wave of severe infections.

Specific questions from the Inquiry team

10.40. I have not answered questions about specific meetings I attended. As I have stated earlier in my witness statement, I attended hundreds of meetings during the pandemic and did not keep personal notes of any of them. I relied on the secretariats of the various organisations who ran these meetings to keep notes.

11. Access to, Sharing and quality of data

Prior to the pandemic

11.1. Planning prior to the COVID-19 pandemic (including the development of the Reasonable Worst-Case Scenario) was informed by analysis of data from prior influenza pandemics. Data sources are given in references to the SPI-M Modelling Summary **{JE/180 - INQ000212124}**. There are only a few pandemics of influenza for which data are available and few interventions were put in place to mitigate the effect of the later pandemics as they were relatively mild. Hence, studies on seasonal influenza or other respiratory infections, such as SARS, were also used to inform this planning document where necessary. In 2019 the World Health Organisation published a series of systematic reviews on the effectiveness of different NPIs **{JE/76 - INQ000212204}**, again with a focus on containing pandemic influenza.

Early period (January and February 2020)

11.2. Cases were concentrated in China during this period, with just occasional cases picked up in travellers from affected areas.

11.3. As explained earlier, data were not readily accessible during this period and were difficult to interpret. Epidemiologists would ideally like access to the "line list", which is

a compiled list of all cases which includes demographic, clinical and epidemiological data on each patient. This is a live database which is updated (usually on a daily basis) to include new patients as well as updated information on the status of each individual patient already on the database (e.g. if they have recovered or died). Line list data can be used to estimate the reproduction number (R number) of the epidemic, as well as delays between onset and report of the case, or onset to hospitalisation and death. Measures of severity, such as the case-fatality ratio can be derived, and how this may change, by different clinical or demographic risk factors. Access to such data is usually restricted for confidentiality reasons. An anonymised version of the database is sometimes made available, though even access to this would usually be restricted due to the possibility of deductive disclosure of patient information. I am not aware that any SAGE participants had access to the Chinese line-list data, nor would I be surprised by this due to the patient confidentiality issues. Various researchers and media outlets tried to compile unofficial line lists from publicly available sources (e.g. **{JE/181 - INQ000212125; JE/182 - INQ000212126; JE/183 - INQ000212127}**). These were used for epidemiological or clinical research in lieu of official, authorised data.

11.4. Aggregate data (e.g. counts of cases, hospitalisations and deaths occurring by day) were more readily available. Count data are not as useful as individual-level data (such as line list data), since aggregate numbers can obscure changes in the underlying dynamics due to delays between infection and the eventual reporting of a case or death, and analysis by subgroup is only possible if the data are stratified in such a way. For instance, comparing the growth rate of the epidemic in two cities is only possible from aggregate data if the data is available separately by city. In the early days of the pandemic the aggregate data from China were particularly difficult to interpret as it was not clear what case definitions were being used (see earlier for explanation). Overall, these problems hampered interpretation of the early data, as noted by SAGE in its minute of February 4th 2020 **{JE/184 - INQ000051925 }.**

11.5. As explained previously in this statement, data on infected travellers from affected regions were also useful. Although few in number, infections in those repatriated was also helpful as they gave an indication of prevalence in the area where these individuals were repatriated from. Chinese data were becoming more available during February (often in the form of peer-review articles) and outbreaks started to occur outside China. That is, data started to become more widely available, partly because

of the spread of the epidemic to other countries and partly because the academic journals started to publish the early investigations of Chinese researchers.

March 2020

11.6. As cases started to accumulate in the UK it became possible to analyse case-based (line list) data to estimate critical parameters such as the UK-specific growth rate and reporting delays. However, there were significant problems with the early data, as outlined in an earlier section. To recap briefly, there were significant delays between onset of illness and reporting of the cases (roughly 6.5 days on average, but this was extremely variable, with some cases taking up to three weeks to be reported **{JE/185 - INQ000212129}**.

11.7. In addition, as most cases being investigated were either travel-related or linked to a traveller **{JE/186 - INQ000212130; JE/187 - INQ000212131; JE/111 - INQ000212043; JE/112 - INQ000212044; JE/113 - INQ000212045; JE/114 - INQ000212046; JE/115 - INQ000212047; JE/116 - INQ000212048; JE/117 - INQ000212049; JE/118 - INQ000212050; JE/119 - INQ000212051; JE/120 - INQ000212052; JE/121 - INQ000212053; JE/122 - INQ000212054; JE/123 - INQ000212055; JE/124 - INQ000212056; JE/125 - INQ000212032}**, the growth rate estimated from such data reflected the growth of the epidemic in travellers and their contacts (and so was reflecting – to an extent - the growth of the epidemic overseas, rather than in the UK). It seemed highly probable that many infections were being missed, given the targeted approach to testing and investigation of cases (we later estimated that between 3 and 10% of cases were detected during March and April 2020) **{JE/188 - INQ000212135}**. Thus, the growth rate and the overall size of the epidemic was obscured by these data problems. These issues could have potentially been avoided by analysing data on the "sporadic" cases, which had presumably been picked up by the recently implemented RCGP and Intensive Care testing system **{JE/111 - INQ000212043; JE/112 - INQ000212044; JE/113 - INQ000212045; JE/114 - INQ000212046; JE/115 - INQ000212047; JE/116 - INQ000212048; JE/117 - INQ000212049; JE/118 - INQ000212050; JE/119 - INQ000212051; JE/120 - INQ000212052; JE/121 - INQ000212053; JE/122 - INQ000212054; JE/123 - INQ000212055; JE/124 - INQ000212056; JE/125 - INQ000212032}**. Given that these "sporadic" cases were not

linked to importations the growth of these should reflect the growth of the UK-specific epidemic. Also, as the sensitivity of both the RCGP and Intensive Care-based system had been roughly calculated, it would be possible to get an estimate of the overall size of the epidemic from these cases. Unfortunately, these "sporadic" cases were not easily identified within the database and were relatively rare due to the modest size of the surveillance schemes. It therefore proved very difficult to estimate the growth rate and size of the epidemic from these "sporadic" cases (see e-mail by Mark Jit for an early attempt at doing this) **{JE/187 - INQ000212131; JE/188 - INQ000212135}**. Added to these difficulties, there were also different versions of the data, with a "line list" a "first-few hundred database" (which was also case-based) and aggregate data starting to be reported on the UK Government dashboard, which had been recently set up. Reconciling these different databases was difficult. That is, trying to match up the numbers of cases for a given day on each of these systems was often a frustrating exercise **{JE/111 - INQ000212043; JE/112 - INQ000212044; JE/113 - INQ000212045; JE/114 - INQ000212046; JE/115 - INQ000212047; JE/116 - INQ000212048; JE/117 - INQ000212049; JE/118 - INQ000212050; JE/119 - INQ000212051; JE/120 - INQ000212052; JE/121 - INQ000212053; JE/122 - INQ000212054; JE/123 - INQ000212055; JE/124 - INQ000212056; JE/125 - INQ000212032}**.

11.8. Overall, the state of the surveillance and data systems was not fit for purpose in the early part of the epidemic in the UK. By that I mean that it was not capable of giving accurate estimates of basic epidemiological indicators, such as the current size of the epidemic and its growth rate.

11.9. The NHS launched the CHESS (COVID-19 Hospitalisations in England Surveillance System) database in mid-March which improved hospital-based surveillance as it was comprehensive and rapid (though count-based, rather than individual level in nature). Detailed analyses of clinical data started to become available from the CoCIN study towards the end of March **{JE/189 - INQ000212136}**. The NHS Sitrep replaced the CHESS data in mid-April. This also contained daily hospital-level count data on COVID-19 admissions and occupancy levels by ward type and age group. The enormous expansion of testing during April 2020 with the data architecture associated with this, greatly helped surveillance efforts. The ONS Coronavirus Infection Survey started to give unparalleled information on infection levels in the community and

household transmission patterns from the end of April 2020. By this time, data availability had been transformed.

11.10. Data access arrangements were ad-hoc at the start of the epidemic. LSHTM signed a data sharing agreement with PHE in early March 2020 and obtained access to individual-level data immediately after this. I do not know what arrangements may or may not have been in place for other institutions and so cannot comment on this. My experience is that once we had made a clear case for access to the data for public health purposes, then PHE acted quickly and professionally to enable this.

11.11. Handling the data in a secure way and making sure that it was clean and in a common format for ease of analyses rapidly became an issue, particularly as the amount and variety of data started to increase in March 2020, some of the problems of the early data are apparent in the emails **{JE/111 - INQ000212043; JE/112 - INQ000212044; JE/113 - INQ000212045; JE/114 - INQ000212046; JE/115 - INQ000212047; JE/116 - INQ000212048; JE/117 - INQ000212049; JE/118 - INQ000212050; JE/119 - INQ000212051; JE/120 - INQ000212052; JE/121 - INQ000212053; JE/122 - INQ000212054; JE/123 - INQ000212055; JE/124 - INQ000212056; JE/125 - INQ000212032}**. In response, the LSHTM CMMID team set up a Data Pipeline subgroup, under the direction of Dr Thibaut Jombart. At the time the data consisted of various forms of aggregated case counts, some anonymised line-lists, RCGP testing, and NHS pathways data. They were all encrypted using libsodium (an encryption software library) and hosted on a private github repository protected by SSH and two-factor authentication. This was updated on a daily basis.

11.12. Collating, cleaning and curating these data on a daily basis was a considerable task. Many of the groups working for SPI-M-O may not have had the resources to keep on top of these tasks. Also, having separate groups undertake this independently was a waste of valuable modelling resources. Hence, during the Spring of 2020 mechanisms were put in place to curate the data centrally for all SPI-M-O members. The analytical team at DSTL (Defence Science and Technology Laboratory) took over the task of collating the aggregate data streams (i.e. counts of new confirmed cases, hospital admissions, occupied beds and ICU admissions and beds and deaths by geographical area), and the SPI-M-O secretariate kept the other databases up-to-date on their secure platform (eXchange). As the epidemic progressed an enormous range of data became available, including behavioural survey data, such as the ONS Opinions and

Lifestyles survey, which included weekly estimates of the uptake of different measures, from handwashing to the booster doses of vaccines and the use of lateral flow tests (we had helped design this survey in the early part of 2020), to the CoMix surveys which gave quantitative information on contacts over time. In addition, travel and educational data, as well as the testing and mortality data were all available. The availability of these data transformed the ability of SPI-M-O to produce analyses, forecasts and projections from late Spring 2020 onwards. There were gaps, however. Data from care-homes, in particular, were not routinely included in the SPI-M-O pack, though counts of outbreaks in different settings (including care-homes) were available from the PHE/UKHSA sitreps.

11.13. The collaboration with data providers was generally good. Representatives from different organisations, including PHE/UKHSA, ONS, the Devolved Administrations, NHS England and front-line staff were present on SPI-M-O. These individuals helped with interpretation of data. Data providers were (or tried to be) responsive to requests for data access from SPI-M-O and other committees. Indeed, from late Spring 2020 the access to data was extra-ordinary – quite literally. It showed what could be possible given the appropriate levels of will and resources. It enabled the epidemic to be tracked to a degree that has never previously been possible and new data to be rapidly analysed and assimilated into scientific advice.

11.14. There are inevitably large gaps in knowledge when it comes to a new disease. These include gaps about routes of transmission; the natural history of the disease, such as the duration of immunity and the degree of protection afforded by infection with one strain to another; or the effectiveness of different clinical or public health interventions. Clinical, laboratory and epidemiological studies need to be designed and implemented (often at scale) during the outbreak to answer these questions. The UK took a leading role in many of these areas, through the establishment of large scale epidemiological studies, including the SIREN cohort study in health care workers **{JE/190 - INQ000212137}** that helped establish - amongst other things - the degree and duration of protection afforded from vaccination and natural infection **{JE/191 - INQ000212138}**; the VIVALDI study in care home workers and residents **{JE/192 - INQ000212139}**; the ONS Coronavirus Infection Survey **{JE/193 - INQ000212140; JE/194 - INQ000212141; JE/195 - INQ000212142; JE/196 - INQ000212143}** and the REACT study that gave unparalleled insight into patterns of transmission in the

community **{JE/197 - INQ000212144}**; as well as the establishment of large scale randomised clinical trials, including the RECOVERY trial **{JE/198 - INQ000212145}**. The analytic capabilities within UKHSA and the ability to link across national databases allowed the rapid analysis of the effectiveness of the vaccination programmes to be evaluated on a regular basis **{JE/199 - INQ000212146}**.

11.15. As data from these different sources - the routine data and the studies undertaken here and overseas - accumulated key uncertainties reduced. Hence, as time progressed knowledge of the natural history of infection and disease as well as the effectiveness of clinical and public health interventions increased, and our uncertainty decreased. As models were continually adapted and fitted to the increasing numbers of data points, the projections became less uncertain. Many uncertainties remained, however, including the degree to which seasonal factors might affect transmission and how behaviour might change as a result of easing or imposing different restrictions. In addition, as new variants emerged, many parameters had to be rapidly re-estimated, including those related to transmission (e.g. basic reproduction number), immunity from vaccines and prior infections, and the effectiveness of further vaccine doses. Thus, considerable uncertainty remained throughout the epidemic.

12. The use of modelling during the COVID-19 pandemic

12.1. A summary of the modelling undertaken by LSHTM CMMID is given earlier in this statement and will not be repeated here. This section will give a brief summary of the process of modelling, and how the models were developed and used to inform decision-making during the COVID-19 pandemic.

12.2. All models developed and used by CMMID (and indeed other groups feeding into SPI-M-O) were bespoke – i.e. developed specifically to inform an aspect of the COVID-19 response and calibrated to specific data streams. There are far too many models and data streams to list them all here, but summaries of the LSHTM models can be found on the CMMID repository **{JE/03 - INQ000092665}** and the Royal Society provides a review of models used to estimate the reproduction number and growth rate **{JE/200 - INQ000212147}**. As explained earlier in this statement, the data streams evolved over the course of the epidemic, allowing the models to be calibrated more accurately. Indeed, at the outset of the epidemic in the UK there were few data to fit to (as there

were few cases, hospitalisations and deaths, and many cases were missed due to a lack of testing). As data accumulated (due to the ongoing epidemic, improvements in surveillance and specially commissioned studies) this allowed the models to be better calibrated, reducing uncertainty. An example is given by the email of Professor Chris Whitty of 30th March 2020 in which he asks what studies would be needed to improve our understanding of the level of asymptomatic infection. The response to Chris Whitty March 2020 **{JE/201 - INQ000212148}** covered planned serological surveys and information on an unpublished study from France. This is one small example of how evidence accumulated over time, helping to inform our view of COVID-19 and how to model it. Estimating the level of asymptomatic infections was important. If many cases were asymptomatic, then for each case reported there would be many more unnoticed and therefore unreported cases. So, one would expect a much higher level of immunity in the population than would be derived from just counting the cases. Unfortunately, the level of asymptomatic infection appeared to be relatively low and by the end of the first wave serological data from PHE suggested that only 5-10% of the population had been infected. The vast majority of us still had no immunity to SARS-CoV-2.

12.3. Different models were developed to answer different questions. The degree of detail included in the model depended on the questions that were being asked and the data that were available to calibrate the model to. So, for instance, no models distinguished between indoor and outdoor contacts, as quantitative estimates of the relative risks of transmission indoors and outdoors were lacking until later in the epidemic **{JE/202 - INQ000212149}**, and SPI-M-O was not asked (as far as I can remember) to model the impact of interventions aimed at just indoors or outdoors contact. It is a common misconception that more detailed models are more accurate. This is not usually the case, as more detailed models involve a greater number of assumptions and require a greater variety of data to inform them (which may not be available). As explained in an earlier section, they are also more difficult to calibrate (fit to data) and the added complexity may mean that the behaviour of the model is difficult to understand. Thus, there is always a balance with model construction to ensure that sufficient detail is included to adequately answer the questions posed, but unnecessary detail is not included.

12.4. At all times SPI-M-O attempted to have at least two independent teams assessing a given question so that the results could be compared against each other (indeed

regular output, such as medium-term projections, were only provided if a minimum of 3 independent models were available). This comparison of model results allows a deeper understanding of any differences that might be found between the models and can give some reassurance of a course of action if different, independent models point to a similar conclusion (see later discussion on the "Roadmap out of lockdown for an example of how this worked"). For routine tasks that were performed on a weekly basis, such as the estimation of the reproduction (R) number, or for short-term forecasting, formal "ensemble modelling" was undertaken. Here, the results of the separate models were combined formally using statistical techniques, to generate an ensemble estimate or forecast. This meant that an overall estimate of (say) the reproduction number and the uncertainty around this could be generated, that included information from each of the separate models (and the data they were fitted to). The methods employed would inherently downweigh any outliers and emphasise the "consensus" view of the different models. Ensemble models generally have better predictive power than the results of any single model. However, they also tend to smooth over differences, and so may be slower to pick-up rapid changes in the true underlying epidemiology.

12.5. As the epidemic developed so the models had to adapt. A good example is related to immunity. There were early signs that immunity to SARS-CoV-2 might not be long-lived (summarised by NERVTAG in 13th March 2020 **{JE/67 - INQ000212195}**). Although estimating the degree and duration of immunity is always important, it turned out not to be critical in the first wave, as only about 5-10% of the population were infected. However, at the beginning of the first wave it was unknown how big it may become (as it was unknown what measures would be adopted to limit its size), so determining the duration of immunity was potentially important (hence my comment to NERVTAG on 13th March indicating the significance of this question at that time **{JE/67 - INQ000212195}**. After lockdown measures were put in place and data on the level of asymptomatic infections became available through the scientific literature it became possible to infer that the size of the wave would likely be relatively small and so the question of immunity became less critical and received relatively little detailed modelling attention until later **{JE/203 - INQ000212150; JE/204 - INQ000212151}** which demonstrates our priorities in mid-March 2020). Once immunity in the population was widespread – largely as a result of vaccination – the degree and duration of immunity against infection, disease and severe disease became critical. Indeed, when

new variants were detected to be spreading, estimating the degree and duration of immunity from prior exposure to previous strains or vaccination was one of the most urgent questions. Routine data could identify re-infection (by linking confirmed cases), but more detail could be obtained from cohort studies. The SIREN study of health care workers **{JE/190 - INQ000212137}** was particularly important in this regard, as it was a large cohort of health care workers who were regularly swabbed (so asymptomatic as well as symptomatic infections could be detected). Given health care workers were more likely to be exposed than others in the population and they were amongst the first to be vaccinated, assessing infection and reinfection patterns in this group would be particularly illuminating. Models were updated to take account of waning immunity, though uncertainty remained regarding immunity parameters and sensitivity analysis was used to explore the impact of different assumptions see for instance, **{JE/171 - INQ000212115}**.

12.6. The above example, and the other Roadmap assessments that the LSHTM team undertook also demonstrate that one of the key uncertainties that remained throughout the epidemic related to the willingness of the public to adhere to lockdown rules, or alternatively, how quickly they might return to pre-pandemic behaviour when measures are lifted. It was possible to measure behaviour via the CoMix survey and from the Google Mobility data. However, predicting future behaviour was not possible to infer with any accuracy. Thus, sensitivity analysis was used to explore the impact of different assumptions on the key outcomes, such as infections **{JE/171 - INQ000212115; JE/172 - INQ000212116; JE/173 - INQ000212117}**. Note that sensitivity analysis is used to illustrate the effect of different assumptions. Note also that predicting future behaviour change is still not possible to accurately do. There are many competing psychological and economic models. However, none have a proven ability to predict the degree of behavioural response to changes in risk. It is interesting that the Danish group claim that their model performed better than UK models of Omicron as a result of including behavioural adaptation, though I have seen no formal evaluation of this statement.

12.7. The above example of modelling to inform the emergence from the third lockdown ("Roadmap out of lockdown") also serves to illustrate the steps taken to try to compare independent model results (to ensure quality) and the difficulties of undertaking peer reviewed research during a fast-moving outbreak. These analyses were undertaken

every month during the stepwise easing of restrictions in the Spring of 2021. There was a 5-week gap between each easement. However, due to the lags in the data (discussed previously) it takes a few weeks for the effect of an easement to start to become apparent – particularly on outcomes that really matter, i.e. hospitalisations and deaths. As the data accumulated after the last easing of restrictions the models were re-calibrated to the emerging data. This gave an up-to-date assessment of the impact of the previous easing (unfortunately, this was complicated by the spread of the even more transmissible Delta virus during this period). The models were then projected forward to assess the possible impact of the future easements and papers were written by each group to describe what was done and what might be expected in the future. The process was then repeated at the next stage (i.e. the next time measures were released). Ideally, each paper would have been subject to external peer review, as this would help ensure the integrity of the findings. Unfortunately, this was not possible under such tight deadlines. Indeed, the LSHTM work was published in the peer-reviewed literature – in August 2022 – more than a year later **{JE/205 - INQ000212153}**, such delays being not atypical for peer-review. Alternative systems for peer review had to be adopted and were, in fact, planned by SPI-M over the many years before the COVID-19 pandemic. The system adopted was to have a number of independent modelling teams assess these critical questions and then compare their results. SPI-M-O organised this process. Three groups were commissioned to undertake the "Roadmap" work: Imperial College, Warwick University and LSHTM. The SPI-M-O secretariat provided an agreed set of parameters relating to vaccination (vaccine efficacy estimates and the speed of the vaccine roll-out) and an agreed set of outputs (e.g. infections, hospitalisations and deaths). Once they received the model projections from the three groups, they compiled and compared them to each other. These were discussed by the SPI-M-O committee and a consensus statement was agreed at each stage (see, for example, **{JE/206 - INQ000212154}**. The presentation of results to SPI-M-O and the development of the consensus statement were particularly important, as differences between the models were explored. These differences were maintained, however, as they reflected uncertainty resulting from different assumptions and data streams used (i.e. model uncertainty). Only if there was an obvious error was a model dropped from a SPI-M-O consensus statement (this did not happen during the "Roadmap", though did occasionally happen for medium-term projections or R estimates).

12.8. The resulting consensus statements were then presented and discussed at SAGE at its next meeting, and then presented to Government. The comparison of the models to each other and the scrutiny of the models and their results by the SPI-M-O committee and then SAGE helped to ensure high standards of scientific output despite the lack of a formal peer review. Finally, it is worth remembering, that within LSHTM we recognised the tension between speed and quality of output. Hence, at the outset of the pandemic, when we set up the CMMID working group, we instigated a system of internal peer review (see earlier). All outputs released onto our repository **{JE/03 - INQ000092665}** had been previously reviewed internally by a panel of CMMID researchers who were not directly involved in the work (later in the epidemic this was also extended to the review of computer code). Although these measures were not as rigorous as external peer-review, they did help us maintain standards whilst responding rapidly to requests and the changing epidemiological situation.

12.9. The systems of review that were put in place during the pandemic (the internal review of LSHTM output and code, and the external review through SPI-M-O) were not as rigorous as full external peer review. Also, there were times, particularly at the start of the epidemic when some groups contributing to SPI-M-O were struggling to catch up with the larger ones in terms of data access, availability of models and human resources. At this time the level of scrutiny that SPI-M-O was supposed to provide may not have been ideal. This was exacerbated by the speed of the epidemic and associated decision-making during March 2020. It is certainly possible that, as Professor Thomas House puts it, "multiple serious errors in models and estimates were presented to Government" **{JE/207 - INQ000056610}**. I cannot remember any serious examples of this, but it is certainly possible. SPI-M-O could, perhaps, examine its processes to try to mitigate this risk. It might make sense for some members to be tasked with reviewing the work of others, primarily (and not expected to contribute new analyses themselves). This might help reduce this risk, though it should be said that such a system could be difficult to maintain as the (academic) incentives for reviewing others' work are largely lacking.

12.10. The necessity to keep models relatively simple so that they could be adequately calibrated, and their behaviour well understood, led the modellers to partition the overall epidemic into different sub-epidemics that were modelled separately. In effect, there were separate models for the community, care home and hospital epidemics, but

no single model that explicitly included all three of these aspects of the epidemic together. Clearly, these sub-epidemics were linked, but the added complexity of a single model did not justify its development (though the Imperial College team did use a model that had care-homes integrated into their "community" model for a while). Indeed, it has been argued that SPI-M-O concentrated too much on community transmission to the detriment of examining the epidemic in specific settings (particularly care homes and hospitals). I think that there is some justification for this view, though care homes and hospitals were not ignored by the modelling groups. PHE/UKHSA led the development of a dynamic model to assess options to limit hospital transmission **{JE/208 - INQ000212155}** later published as **{JE/209 - INQ000212156}**. Within CMMID, Dr Sam Abbot developed and applied methods to track the reproduction number for the hospital epidemic **{JE/210 - INQ000212157}** and Dr Gwen Knight led a consortium to quantify the burden of health care associated infections **{JE/211 - INQ000212158}** later published in the peer-review literature as **{JE/144 - INQ000212082}**. Alicia Rosello of CMMID led the development of a model to assess measures to limit the occurrence and severity of outbreaks in care homes **{JE/212 - INQ000212159}** later peer reviewed and published as **{JE/213 - INQ000212160}** and Ian Hall (University of Manchester) chaired a SAGE working group on care homes. Later in the epidemic, the LSHTM team also looked at vaccination strategies for prisons **{JE/214 - INQ000212161}**. The studies into the care homes and hospital outbreaks were regularly discussed at SPI-M-O and papers were provided to SAGE **{JE/215 - INQ000212162}** Thus, these epidemics were not ignored by the modelling community, but further analyses would have been very welcome, given the importance of trying to prevent cases occurring in these settings and SPI-M-O's aim to have multiple, independent modelling input for key questions. Modelling resources were stretched very thinly during the pandemic. Better co-ordination and prioritisation could have led to more equal coverage of all of the relevant modelling questions. It is, however, difficult for SPI-M-O to ensure this. SPI-M-O commissioned (asked for) work but had no budget to pay for it.

12.11. The most common model structure employed for population-based policy decisions were variations on age-structured SEIR (susceptible, exposed, infectious, recovered) models. In these models contact within and between age groups is governed by an Who-Acquires-Infection-From-Whom (WAIFW) matrix, the elements of which can be

estimated from contact studies, such as the POLYMOD study from before the pandemic **{JE/45 - INQ000092682}** or the weekly CoMix estimates derived throughout the pandemic **{JE/60 - INQ000212188}**. Note that no other structures, such as schools, households, care-homes or hospitals are explicitly represented in these models. Instead, the WAIFW describes the mean contact rates across different age groups. Social distance measures, such as shielding of the elderly or working from home mandates can be implemented in the model by adjusting these contact rates (e.g. **{JE/59 - INQ000212187}**). These age structured models were also adapted to take into account the age-specific risk of disease, hospitalisation and death following SARS-CoV-2 infection (e.g. **{JE/59 - INQ000212187}**). These age-specific rates had been relatively well established by early March 2020 **{JE/13 - INQ000092647; JE/142 – INQ000212080; JE/216 - INQ000212164; JE/216A – INQ000255452}**), and all the models used to assess population-based policy options included this dramatic increase in risk in elderly individuals (e.g. **{JE/59 - INQ000212187}**). Note that this simple model structure does limit the policies that can be evaluated by this population-based model. Policies aimed at reducing transmission to or within care-homes, for instance, are better evaluated using a specific model designed to address these issues specifically (e.g. **{JE/212 - INQ000212159}**). As mentioned earlier, with the exception of the teams at LSHTM and University of Manchester, there were few attempts to model transmission in care homes, possibly because of a concentration of effort on population-based models.

12.12. It would be wrong, however, to give the impression that these age-structured SEIR models were the only models used for policy purposes. There were dozens of bespoke models (possibly hundreds) developed to answer specific policy questions, from testing and quarantine strategies at the border (e.g. **{JE/07 - INQ000092696; JE07A – INQ000255395; JE07B – INQ000255396; JE07C – INQ000255397; JE07D – INQ000255398; JE/217 - INQ000212165}** to the effect of social bubbles **{JE/218 - INQ000212166}** and contact tracing and isolation **{JE/11 - INQ000092645; JE11A – INQ000255406; {JE/22 - INQ000092657; JE/22A – INQ000255418; JE/219 - INQ000212167}**. Contrary to claims by Mr Jeremy Hunt **{JE/220 - INQ000212168}**, considerable work was undertaken by LSHTM in the very early stages of the epidemic to evaluate the effectiveness of contact tracing and isolation, with a paper online on 7th February 2020 **{JE/11 - INQ000092645; JE11A – INQ000255406}** that was discussed

by SPI-M-O and published in a peer-review journal by the end of that month. Indeed, this early paper was extremely influential, having been cited in the literature over 2700 times at the time of writing (Google Scholar).

12.13. The co-chairs of SPI-M-O (Professor Graham Medley and Dame Angela McLean) along with the SPI-M-O secretariat handled requests for modelling work from Central Government (mainly the Cabinet Office). As far as I am aware, they devoted considerable time to ensure that the limitations of the models were understood by policymakers and that the questions that were asked were potentially answerable. Officials were also present at SAGE and SPI-M-O meetings. The co-ordination with Government improved over the course of the epidemic. In particular, the engagement of Cabinet Office in the modelling process seemed to increase significantly with the appointment of Rob Harrison in late 2020, which undoubtably helped improve understanding of the appropriate use and potential limitations of modelling. This was also facilitated by the secondment of modellers to the Cabinet Office (e.g. Dr Nick Davies from LSHTM) and specific training ("Teach-in") sessions that were run to help improve understanding. The Developed Administrations also had their own technical advisory groups that were attended by senior modellers (e.g. Professor Medley) when required. To what extent senior politicians knew about the strengths and limitations of modelling is not known, though I believe that Mr Sunak (then the Chancellor) attended one of the teach-in sessions on modelling that I ran on the 1st of June 2020. The session was arranged to allow Mr Sunak to attend. I cannot confirm whether he attended or not as the session was run on zoom.

12.14. Despite these efforts, there were times when modelling analyses were handled very badly. Two obvious examples relate to the provision of model estimates for the Reasonable Worst-Case scenarios in the summer and autumn of 2020 and the misuse of the latter to justify the second lockdown. In July, the Cabinet Office asked SPI-M-O to update model estimates for the Reasonable Worst Case-Scenario **{JE/221 - INQ000212169}**. This request consisted of a series of arbitrary assumptions about incidence doubling at specific points in time, after which it is held constant for a given period as (unspecified) measures come in place, etc. There was no epidemiological justification for any of these assumptions. It seemed to many of us that it was over-specified and arbitrary and in no way represented a reasonable worst-case scenario. That is, it was very easy to think of scenarios that were much worse than this which

had a reasonable chance of coming into being. The assumptions were also so simple that they could have been worked out on a spreadsheet (double incidence at this time, double it again, etc). There was no need to spend valuable and sophisticated modelling resources on this. Nevertheless, the different SPI-M-O groups did produce these scenarios for the Cabinet Office. As expected, the real epidemic started to overtake these Reasonable Worst-Case scenarios in September 2020. At this point the SPI-M-O Secretariat started the process of coming up with new Reasonable Worst-Case Scenarios. I assume that they wanted to "own" the process this time, given how poorly the previous ones had been constructed and performed. However, the initial attempt to come up with updated Reasonable Worst Case Scenarios ended up being used as justification for entering the second lockdown at the end of October 2020 which damaged the credibility of the modelling process at a particularly sensitive and politically highly charged time **{JE/222 - INQ000212171; JE/223 - INQ000212172}**. These initial, updated scenarios had been developed by four independent SPI-M-O groups (Imperial, Warwick, LSHTM and PHE/Cambridge) at the start of October in response to a commission from SPI-M-O. As the name implies these were reasonable worst-case scenarios. They were not intended to be an expectation of what was most likely, but an upper bound of what was likely. In addition, these were draft scenarios – not completed pieces of work – and they were intended to inform an overall single Reasonable Worst-Case Scenario to be negotiated with Cabinet Office. Somehow these were used to inform government policy and – in some parts of the media at least – were treated as predictions. Particular political and media attention was paid to the most pessimistic of these four different reasonable worst-case scenarios (produced by PHE/Cambridge) which stimulated headlines of 4000 deaths per day being expected. The modelling teams (and modelling in general) were then heavily criticised in the press and by some politicians when reality did not match these "predictions", with the resulting reputational damage that this entailed. What is worse is that SPI-M-O were producing weekly "Medium Term Projections" at the time. These used the same four models to produce an expectation of what might occur (instead of an upper bound). These Medium-Term Projections combined the individual projections from the four models to form an "ensemble" projection that therefore down-weighed extreme projections (i.e. the more pessimistic or optimistic ones). These were being produced every week and the projections were being compared to the subsequent data. That is, the Government not only had more rigorous projections available to it, but it also had

information on how well these projections were performing at the time. To highlight the most pessimistic aspect of an out-of-date, draft, reasonable worst-case scenario to justify their decision for the second lockdown was clearly a mistake which damaged the reputation of the modelling teams, the modelling process and the Government **{JE/224 - INQ000212170}**. These examples serve to show the difficulties inherent in trying to use models appropriately to inform Government planning. That examples such as these were relatively rare is testament to the work of the SPI-M-O Secretariat and its Co-Chairs in ensuring that modelling was used appropriately by central government.

13. The "R" number

13.1. The R (or reproduction) number is the average number of secondary cases a typical case generates. It is closely tied to the growth rate of the epidemic as if the R number is greater than 1 then each case is (on average) causing more than one more case and the epidemic will be increasing (the growth rate will be positive). If the R number is less than one, then the epidemic will be shrinking in size (as each case, on average, is generating fewer than 1 more case). The R number is sometimes more formally called the effective or net reproduction number. It is also closely related to the Basic Reproduction Number (R_0), which is the reproduction number when the entire population is susceptible, and no control measures are in place (which was the case for SARS-CoV-2 infections in early March 2020). The basic reproduction number gives a maximum value for the R number. This was estimated to be about 3 in March 2020. To control an epidemic the reproduction number needs to be maintained at or below 1. That is, it had to be reduced by about 2/3rds in March 2020 to stop the epidemic from growing further and stop the NHS from being overwhelmed.

13.2. In a simple, homogenously mixed case (which is an oversimplification), then the reproduction number, *R*, can be written as:

$$R = p\ c\ D\ s$$

where *p* is the probability of infection, given contact, *c* is the contact rate, *D* is the duration of infectiousness and *s* is the fraction of the population who are susceptible. At the outset of the epidemic, virtually everyone was susceptible to SARS-CoV-2 (there was no immunity in the population) and so *s* equalled 1 (in the equation above).

13.3. Measures to reduce the likelihood of infection, given contact, affect p. These might include mask wearing, physical barriers, hand hygiene and maintaining social distance (e.g. the 2 metre "rule"). Measures to reduce c include closing schools, workplaces, bars and restaurants and stay at home orders (lockdown). Measures to reduce the duration of infectiousness include effective isolation (removing infectious cases). If we can estimate the contribution that different measures make to the overall R number, then it should be possible to design a package of interventions that would maintain R at or below 1 (perhaps also picking those that would minimise the impact on the economy or some other aspect of society). In practice, this proved very difficult. Measures were put in place simultaneously making it very difficult to estimate the impact of any single intervention. Similarly, when measures were lifted, groups of measures were eased with little time between the easing steps to assess the impact of lifting any one group of measures. International comparisons could help, but the differences in interventions and how they were implemented, as well as inherent differences in populations, contact patterns and surveillance systems also made this very difficult.

13.4. There are many different ways that the reproduction number can be calculated. The most common methods calculate the reproduction number from the growth rate of the epidemic, with knowledge of the generation time distribution (the time between generations of infections) **{JE/225 – INQ000212173}**. These methods analyse a time series of data – typically reported cases, hospitalisations or deaths. However, each of these data streams are subject to different delays from the point of infection. Data on deaths are more lagged than data on cases, so estimates of the reproduction number based on death data will reflect infections that occurred further back in time than estimates based on reported cases.

13.5. A number of groups contributing to SPI-M-O estimated the R number on a weekly basis. The different groups used a variety of methods and concentrated on different data streams (e.g. some fit to the case numbers, others to hospitalisations, etc). These were combined into an ensemble (or consensus) estimate of the reproduction number on a weekly basis. As all of the data streams were lagged (with respect to infection) and it took about a week from the point when the last data was included for the signed-off estimates of R to be published, the published R number was actually reflecting the epidemiological situation some weeks before. The ensemble methodology also tended

to smooth over variations (e.g. a sharp change in the reproduction number might be picked up relatively quickly in the least lagged data stream, but not the others, so by averaging over all the data streams this sharp change would be flattened). Overall, the published R number was not very sensitive to changes in the epidemiology, which limited its usefulness.

13.6. Furthermore, the routine published R estimates were calculated at a course spatial scale (the countries within the UK and English regions). This might mask considerable variation at a lower spatial scale. For instance, there might have been hotspots of increasing incidence in a region of declining incidence (the weekly SPI-M-O spatial subgroup did track and report on the reproduction number and growth rates at finer spatial scales). Hotspots could also have occurred in specific settings. For instance, in April 2020 the overall epidemic was declining (R was less than 1) but outbreaks were occurring in many care homes across the country. Although cases occurring in these outbreaks would have been recorded in the routine data streams and therefore contributed to the overall estimate of R, as they are relatively rare (in comparison to the entire population) they would have a relatively small impact on the overall estimate of R. That is, a policy that keeps the reproduction number low in society in general might well have to be supplemented by additional measures in higher risk groups.

13.7. In summary, the reproduction number can be a useful indicator. However, it should not be used alone to guide policy decisions, not only because the overall number is relatively crude (masking temporal, spatial or social heterogeneities), but also because there are very different implications of maintaining R at around one when incidence is high versus low. Keeping the reproduction number close to one means keeping the incidence roughly constant. However, maintaining a high incidence will result in far more hospitalisations and deaths over time than keeping the reproduction number at or around one when the incidence is low. If the policy goal is to “save lives and protect the NHS” then maintaining a low reproduction number (i.e. R less than 1) across all areas of society can help achieve this. However, maintaining the reproduction number at or near one may not (it depends what the incidence is, when it is kept constant). Letting the reproduction number increase to be above one across society will certainly not help “save lives and protect the NHS”.

14. The influence of modelling on policy decisions during the COVID-19 pandemic

14.1. How many cases might there be in two weeks' time? How many deaths? How effective will closing schools be at reducing incidence? What if we also ask all non-essential workers to stay at home? All of these questions, and many others like them, require a model to answer them. The future has not happened yet, and so it is not possible to observe what will happen. Some means of being able to project our best guess about what might happen under different scenarios is therefore needed. Indeed, if an explicit mathematical model is not available to provide projections to help answer these questions, then the only recourse would be to use opinion (expert or not). Ironically, an expert attempting to answer such questions would need to construct a mental model of the epidemic. Making this model explicit, stating its assumptions, calibrating it to data sources (and wherever possible validating its projections against external data sources), is a far more open, accessible, and rigorous process.

14.2. Indeed, the same can be said for social or economic projections and scenarios. What would the short and long-term costs of school closure be? What would be the social and economic impact of instructing non-essential workers to stay at home? Again, these questions can only be answered by a model – either an implicit one (in an expert or policy-maker's head) or an explicit one, that can be challenged and rigorously tested against data.

14.3. Hence models are an essential tool for policymakers. The primary problem in the use of models to inform policy during the pandemic was not that there were too many epidemic models (or modellers), but that there were not enough assessments (including model-based assessments) of the economic and social impact of the different policies. Economic assessments that might have been produced were generally not open to scrutiny by peers, the press, and the public, so it is impossible to tell whether they were well founded and could explain the subsequent data. This Inquiry is examining the epidemic modelling work in intimate detail. This is partly because these models were explicit, and they were made available for scrutiny. Is the Inquiry spending as much time on the economic models? If not, why not?

14.4. Of course, there are limitations to any modelling work, and these weaknesses need to be clear to policymakers. SPI-M-O attempted to spell out the limitations of the modelling work, through its consensus statements as well as through maintaining

dialogue (via the Chairs and secretariat) with policymakers. An example of an attempt by SPI-M-O to make clear the limitations of the work is given in the last point of the March 20th 2020 consensus statement (though the previous 9 points in the consensus statement give a clear assessment of the epidemic at the time, its approximate rate of growth, its likely impact on ICU admissions in London and elsewhere and how quickly interventions might lead to a reduction in pressure on intensive care) **{JE/226 - INQ000228591}**.

14.5. As I have explained in earlier sections, although epidemic models have a strong theoretical foundation and a long history of use to inform policy-making, they are relatively crude tools. They are a simplification of incredibly complex systems, such as our society and how our interactions might lead to the spread of infections and how our actions along with our innate and adaptive immune responses might limit their spread. Despite this simplicity, they can be used to give qualitative and even quantitative insights as to future epidemic trajectories under different assumptions. However, as with all tools, their limitations and weaknesses need to be appreciated along with their strengths. Whilst I think that there were relatively successful attempts to improve the understanding of the use of models within government (for example by secondments of academic modellers into Government (including the Cabinet Office) or Government agencies, such as the Joint Biosecurity Centre and PHE/UKHSA), there were deficiencies in understanding that persisted throughout the pandemic. As mentioned elsewhere, a common misunderstanding, often propagated in elements of the press or amongst some politicians, was the distinction between model-based scenarios and forecasts. Scenarios allow policymakers to assess the possible consequences of different courses of action – what might happen if policy A is adopted, or policy B. These are not forecasts, which are short-term predictions of what is likely to happen, for the simple reason that the decision-maker chooses a policy-option and therefore determines (to some extent) which future epidemic trajectory occurs. This does, of course, result in a problem as the alternative course of action did not happen, so the only way to assess what might have happened (i.e. the counterfactual) is to rely on a model. We did not experience hundreds of thousands of deaths in the first wave of COVID-19 (though over 40,000 did die), because the government took action to avoid this possibility (we entered lockdown). However, the only way to assess how many deaths might have been avoided by this course of action is to use a model that

can project the consequences of alternative scenarios (e.g. **{JE/138 - INQ000212078}**). By confusing, deliberately or otherwise, scenarios with forecasts, critics can always point to the scenarios describing policy options that were not adopted (e.g. the do-nothing scenario during the first wave) and say that it didn't happen and therefore the models were wrong. Or even worse that the modellers were trying to "create a climate of manipulative fear". Given models are used across wide areas of Government to help evaluate different scenarios, improving the dialogue around their use would greatly improve public and political understanding. Indeed, improving the level of scientific understanding more generally as well as the quantitative abilities of our civil service, government, and elected representatives would stand us in better stead for the next crisis and almost certainly improve routine policymaking as well.

14.6. Note that models sometimes have to make assumptions because there are no data to inform them. Modellers try to avoid this (which is one reason why simple models are often preferred to more complex ones) but at times it is inevitable. This is particularly true at the outset of a new pandemic when there are many uncertainties and many unknowns. As data accumulates these uncertainties tend to reduce and parameter estimates can be made or models can be fitted (calibrated) to emerging data, reducing the range that unknown parameters can take. A good example of this is provided by the first national lockdown. What fraction of individuals would comply with these measures? That was unknowable at the time as we had never gone into lockdown before, neither had any of our neighbouring countries. Thus, assumptions had to be made. Although these assumptions were unsupported by any evidence at the time, they were at least explicit and could be challenged, as well as be subjected to sensitivity analysis. As data on compliance with measures started to accumulate (via ONS surveys and CoMix, for example) these assumptions could be revisited and more accurately parameterised. That is, later projections would have reduced uncertainty surrounding this aspect. Would it have been better if the epidemiological community had not assumed anything and refused to offer any explicit modelling projections on the possible impact of lockdown? Policymakers would still have wanted to know what the impact of lockdown might be. This void would then have been filled by implicit models, that can't be challenged or examined.

14.7. It was always clear that a severe pandemic, such as the one that we have just experienced, will impact all aspects of society and could have major economic effects, both as a direct consequence of high rates of disease and death in the population and as a result of the measures that may be put in place to limit its spread (e.g. **{JE/227 - INQ000212174}**). It is clear, therefore, that decision-makers need to weigh the potential public health impact of a course of action with the potential societal and economic costs. Committees and procedures had been put in place long before the pandemic so that scientific evidence (including epidemic projections) could be collected, analysed, assessed, and then presented to decision-makers (these included, SPI-M, NERVTAG and SAGE along with the Joint Committee on Vaccines, and Immunisation, **('JCVI')**. These structures were expanded during the pandemic through the standing up of several other committees (such as SPI-B and many subcommittees of SAGE) and the setting up of new Government agencies (such as JBC) as well as the expansion of existing ones (such as PHE/UKHSA). Indeed, other government agencies or departments also set up further scientific committees, such as the UKHSA's Testing Initiatives Evaluation Board and its Variants Technical Group. These initiatives and committees, along with the enormous expansion of our surveillance systems and the specialised scientific studies that were set up ensured, on the whole, that policymakers were availed of the most up-to-date assessment of the progress of the pandemic as well as the potential clinical and public health impact of different policies. As far as I am aware, the scientific information that was assessed in this way was publicly available and so open to scrutiny and debate.

14.8. It was not clear to me, however, to what extent, if any, the government had set up similar structures to assess the economic impact of the pandemic and of different policies to control it. No equivalent of SAGE was set up to assess the evidence and its quality, and very few, if any, economic analyses were in the public domain. If such analyses were performed, then they were not open to the same oversight as the scientific evidence. This imbalance in the weight and quality of the evidence on the scientific and public health side versus the economic side, seemed to me, to risk poor decision-making . The solution was clearly not to be less well informed scientifically, but to be equally well-informed on the economic aspects of the pandemic.

14.9. Given the enormous economic impact of the pandemic and that this was well known to be likely before the pandemic occurred (hence pandemic influenza's place on the

national risk register) it is surprising that no such measures seemed to be taken over the two and a half years of the acute phase of the pandemic. Every policy that the Government adopted, or considered and decided not to adopt, would have led to potential economic costs and potential economic benefits. However, there seemed to be no formal arrangement to draw in the enormous economic expertise in academic and other institutions across the country to assess and evaluate the strength of economic evidence underpinning a possible course of action (including the use of economic models). Nor was there any attempt to link epidemic modellers with economists to try to formulate a more holistic assessment of the impact of alternative courses of action. It is hard to believe that this failure of omission did not lead to poor decision-making. How much external scrutiny was "Eat Out to Help Out" subject to before it was launched? Was the delayed decision to enter lockdown in the autumn of 2020 influenced by economic analyses? I don't know the answer to these questions as almost nothing is in the public domain – in stark contrast to the scientific advice and the evidence in support of this.

14.10. SAGE's remit was limited to offering scientific advice. It did not extend to economic analyses and was not constituted to take this on. This was clear from the outset and members were reminded of the limitations of SAGE on a periodic basis. As the epidemic progressed, SAGE membership expanded to cover other areas of science, including social sciences. It would have been very difficult, and in my view inadvisable, for SAGE to take on another major area of advice, i.e. offering economic advice. However, I do not see why a parallel system could not have been established.

14.11. Had economic data and models been analysed and scrutinised to the same extent that epidemic models were, this might have led to a more open and honest debate about what we, as a society, wanted to achieve and how we should best go about it.

15. Transparency and communication of scientific advice

15.1. Transparency is critical in maintaining trust in policy decisions. Secrecy leads to speculation which can easily lead to distrust in the process of policy formulation and mistrust in the aims and intentions of the Government. Trust is critical in maintaining compliance in any public health intervention, where individuals are typically asked to take actions to reduce their risk of disease or injury – actions that often limit their freedoms to some extent. This was particularly true during the pandemic as the

Government were putting in place unprecedented measures that affected everyone's liberties and potentially threatened their livelihoods and wellbeing. Expecting individuals to undertake these measures without explaining in detail why they were necessary would be a significant risk. Transparency can also improve the standard of debate and the standard of the evidence itself, as it will be scrutinised. It seems obvious that the Government were slow to adopt transparency, and even then, only did so patchily, as mentioned in the previous section the openness of the later SAGE advice and evidence underpinning it contrasts with the economic advice the Government may have received. Initially, SAGE membership and papers were not publicly available. There were various reasons put forward for this including the personal security of SAGE members and possible, unspecified, issues of national security. Most SAGE members that I spoke with at the time thought that these concerns were outweighed by the benefits that would arise from openness. Indeed, speculation on SAGE membership was becoming a distraction in the spring of 2020, as was the call for openness – a call which many of us on SAGE agreed with. This secrecy also led to others stepping into the public debate about different policy options, including Independent SAGE, who held their meetings in public. As it says on Independent SAGE's website, "We believe openness and transparency leads to better understanding and better decision making. We also believe it the responsibility of scientists and those with specialist knowledge to engage with the public and policy makers, in order to ensure that science benefits all of society." **{JE/228 - INQ000212175}**. Many of us on SAGE shared these views and were relieved when the membership and papers became publicly available.

15.2. Following this change, SAGE papers would generally be released within a week or so of the SAGE meeting – the delay allowing the Government to discuss any policy implications in private. This seemed reasonable to me. I only remember one episode when the delay to publication seemed to be excessive, which was around the decision to not adopt the package of measures that SAGE recommended in September 2020 to mitigate the autumn wave. The Government clearly did not want to heed this advice and took some weeks to release these papers – which helped build speculation and anticipation. If the Government had been open as to the scientific advice, any contrary advice (e.g., economic impact assessments), and about its overall aims, then it could have explained why this decision was made.

15.3. This example also serves to illustrate the vacuousness of the "following the science" phrase that government ministers frequently repeated, particularly in the early part of the epidemic. The Government's job was to weigh scientific, economic, and social impact assessments and to come to a decision based on all of the evidence, the strength of this evidence (or otherwise), their overall strategic aims, and the prevailing operational constraints. The scientific evidence was only ever a part of this, and was frequently very uncertain, particularly at the outset of the epidemic or when new variants arose. The phrase was therefore inaccurate on two counts:

15.3.1. there was no single "science"; and

15.3.2. the Government did not and should not follow scientific advice to the exclusion of all other evidence and operational constraints.

This phrase was also damaging as it "blurred the line between scientific advice and policy decisions" as the Institute for Government put it. Many of us felt that ministers – i.e. decision-makers – were using this phrase to hide their role in making unpalatable decisions. A more accurate phrase would have been that minsters "took account of the scientific evidence". Openness over the role of advisers and decision-makers, the strategic aims of the Government, the constraints that it was operating under, and the role of evidence in these decisions could have helped improve trust in the Government, its advisory bodies, and the policies themselves.

15.4. One of the central problems throughout the pandemic was the lack of a clear strategic direction from the Government. This made giving advice more difficult, as it was not clear what the Government wanted to achieve (crudely put, was the aim to minimise deaths or minimise the impact on the economy). It seems likely that this led to delayed and inconsistent decision-making. This was particularly apparent in the first year of the epidemic, when a suppression strategy was adopted in mid-March 2020 after a mitigation strategy had been announced in early March. This was replaced, in late Spring and Summer 2020 with encouragement to return to work, encouragement to travel abroad and even public subsidies to take epidemiological risks (i.e. to eat within restaurants).

15.5. A clear and consistent strategy would also have helped the Government to adopt clear and consistent messaging. Instead, the flip-flopping between economic and health priorities led to confusing public health messaging (brilliantly summarised by the

comedian Matt Lucas in May 2020 **{JE/229 - INQ000216667}**. Indeed, the inconsistent messaging may well have undermined public health communications and could have endangered health. I remember being particularly alarmed by the Prime Minister – before the first lockdown – suggesting that the epidemic will have turned a corner in 12 weeks (e.g. **{JE/230 - INQ000212176; JE/231 - INQ000212177}**). Confusion in public messaging can undermine confidence and reduce compliance or result in important messages being lost. Again, a clear strategy with the evidence underpinning it (even if there was no evidence and the approach was a precautionary one) could well have helped improve messaging and build trust in the policies that were being adopted and the advice that was being offered.

16. Lessons Learned

Strategic issues

16.1. The Government's aims were unclear, and it undertook multiple reversals of strategic direction over the course of the epidemic. It seems very likely that this contributed to poor health and economic outcomes. The first national lockdown was effective at reducing incidence. It drastically reduced hospitalisations and deaths, reduced the risk of transmission to vulnerable people in care homes and elsewhere and broke up chains of transmission giving the Test and Trace system a better chance to operate effectively. All this came at enormous social and economic costs. Instead of capitalising on this hard-won strategic position, it was wasted by a raft of measures in the summer of 2020, including - most perversely - encouragement for risk taking through government subsidy ("Eat Out to Help Out"). This short-term economic sugar-rush inevitably resulted in another crash. Unfortunately, when the second lockdown came it was insufficient to return the incidence to low levels. We entered the crucial winter period with cases and deaths high and the NHS already under severe strain. Along with poorer health and economic outcomes, this lack of a long-term vision and plan (particularly in the first year) led to inconsistent public health messaging and complicated the role of advisors, scientific or otherwise. It is difficult to plan a course when the destination is unclear.

16.2. Epidemics grow exponentially. Even the best public health surveillance systems will lag behind the epidemic as it takes time for disease to become apparent and cases

will always be missed. Even a few days delay can be crucial if the epidemic is doubling every few days. This means that decisions need to be made quickly and waiting for better information may be very costly. It also means that measures need to be put in place that appear to be wider and harsher than is necessary, as the epidemic will be more widespread than is apparent from the data. In circumstances such as this, where delays can be very costly, there is a need for the precautionary principle to apply. SAGE's advice attempted to be neutral (particularly at the outset of the epidemic), laying out the scientific evidence in a balanced way (e.g. on the effectiveness of travel restrictions and the lack of clear evidence on mask wearing). **{JE/232 - INQ000212178}** contains a list of the early SAGE recommendations in the first few weeks of the pandemic (January and February 2020). Apart from the first of these, on entrance screening, they are not really recommendations at all. For instance, the statement in SAGE's minutes of the 27th of February that "Modelling suggests earlier and/or combined interventions will have more significant impact" **{JE/233 - INQ000106129}** implies that a package of interventions should be introduced quickly but falls short of actually recommending it. At times SAGE even anticipated that certain measures would be unpalatable or difficult to implement. Added to this, the Government repeatedly showed significant optimism bias **{JE/231 - INQ000212177}**, meaning that they tended to overestimate the likelihood of positive outcomes and underestimate the likelihood of negative outcomes. Some examples of where this occurred are **{JE/234 - INQ000228593; JE/235 - INQ000251522; JE/236 - INQ000228594; JE/237 - INQ000232351; JE/238 INQ000228596}**. The neutral (scientific) tone of the advice from SAGE met with this optimistic outlook, which may well have led to significant delays to effective action taking place. Later, SAGE shifted in tone to become more precautionary, most notably in September 2020, though this was not effective at bringing about more rapid action. Multiple international comparisons suggest that a precautionary approach could lead to better health and economic outcomes and it should underpin policy-making in this area (e.g. **{JE/239 - INQ000212180}**).

16.3. NPI's were a stop-gap before a permanent solution to the epidemic could be found – that permanent solution being high levels of immunity (ideally through vaccination) so that other measures would no longer be necessary. Widespread vaccination was likely to take one to two years (at least) to achieve. The build-up of natural immunity would

likely take a similar amount of time if the NHS was not to be overwhelmed in the process. This had important implications. NPIs were always going to have to be in place for a significant period of time. Businesses, organisations, and the public needed to be aware of this so that they could plan appropriately. This did not happen. Avoiding hard truths does not make them go away but can increase their impact if society is not prepared for them. Secondly, the time taken to achieve herd immunity through natural infection or vaccination was likely to be of the same order of magnitude. In addition, the ongoing costs of low incidence (vaccine induced immunity strategy) or high incidence (natural immunity strategy) was likely to be similar as both involved holding R at around one. However, one of these scenarios would result in high numbers of deaths and pressure on the health service, whereas the other would not. The failure to grasp these strategic issues carried an enormous cost. Epidemics are fast moving and require rapid responses. Despite the immediate pressures, there remains a need to think carefully over the longer term to help set an overall approach to managing the crisis.

16.4. Improving the standard of scientific literacy is essential in our public life.

Planning and preparedness

16.5. The international alarm about what became known as SARS-CoV-2 and its associated disease (COVID-19) was raised at the end of December 2019. By late-January 2020 it was clear that this was likely to spread rapidly around the world, if it hadn't already, and that we were probably in the early phases of a pandemic. It was also clear that this virus was pathogenic as well as being transmissible and so large numbers of deaths should be expected, indeed, in terms of possible number of cases and deaths it looked similar to the 'Reasonable Worst Case Planning' assumptions at the time, which were largely based on the influenza pandemic of 1918. We had a period of around 6 weeks when actions could have been taken to mitigate the effect of the epidemic. As this pandemic would likely affect every aspect of society there would be major implications for every government department and every aspect of society, not just the health and care sector. I was not party to the preparations that were being undertaken within Government, so these comments must be taken in that light – i.e. the view of an outsider. However, I could not see that much was being done. When

the epidemic really took hold in mid-March and measures were put in place to limit its spread, these seemed to come as a complete surprise to most people, businesses, and institutions. Society was not prepared. Across every sector, measures were hurriedly put in place, often in an ad-hoc way. My impression is that the Government did not really take the pandemic very seriously until March, by which time it was very late. Much of the lead time had been wasted.

SAGE and other scientific advisory bodies

16.6. SAGE did not set overall strategy – this was the Government's role. However, one area where SAGE could have been used more is in the framing of realistic scenarios that could be used to inform strategy choices. From the epidemiological, scientific and public health side, SAGE was well placed to perform forward looks, but seldom did so (the 'Roadmap out of Lockdown' was the closest to this **{JE/36 - INQ000092672; JE/62 - INQ000212190; JE/171 - INQ000212115; JE/172 - INQ000212116; JE/173 - INQ000212117; JE/208 - INQ000212155}**). Presumably, these sorts of scenarios were being worked on elsewhere in Government (at least I hope they were). Whilst I think that the separation between advice and strategic and operational planning is important, I feel that the expertise on SAGE could have been better employed to help shape strategy more than it was.

16.7. As stated above, a precautionary approach is needed when dealing with a fast-moving epidemic. Whether this approach feeds in at the stage of the scientific advice or at the stage of policy formulation, is perhaps a smaller point, but one that needs to be clarified. It was not clear to me whether SAGE advisors were expected to apply the precautionary principle when giving advice or not. Adopting a view about risk and precautionary measures impinges on the domain of policy- or decision-makers. Furthermore, by applying the precautionary principle, policies may be recommended without proof of their scientific effectiveness. This could lead to legitimate challenge from sectors or individuals adversely affected by such a policy, thereby undermining advice more generally. A clear stance on how advisors should act with regards the precautionary principle (right across government) would be helpful.

16.8. It was similarly unclear what the operational, ethical, and strategic constraints were, further complicating advice. The Government presumably had some level of deaths

that was deemed unacceptable (and by implication a range that was acceptable) but it was never clear what this might be. It was also unclear (at least to me) what interventions were unacceptable as they were likely to be too costly or unpalatable. This may well have led to SAGE advice being too timid, particularly in the early days of the epidemic (that is, we did not anticipate that more radical measures, such as lockdowns were potentially on the table and so did not spend much time on them). Likewise, the constraints on the health service were never clear. Obviously, there is considerable flex in any system as large as the NHS and so there is not a hard limit on the availability of beds and intensive care facilities. However, a better indication of the tolerable range would have helped. Perhaps it is naïve to expect that the political, ethical, and operational constraints would be spelt out (and they almost certainly changed over time). Nevertheless, clarity in this regard could have helped tailor advice and focus effort.

16.9. The lack of information on constraints partly resulted from the way SAGE was organised to feed into the decision-making process. Contact with decision-makers occurred through CMO/CSA. There was little or no contact outside this. This isolated the scientific advisors from the decision-makers, which enabled scientists to engage in technical conversations and helped maintain the integrity of the scientific advice as it was free from external interference. However, as stated above, it was difficult to gauge what was feasible. It was also not clear whether decision- or policymakers- adequately understood the evidence and its limitations. Politicians and senior civil servants did not have the opportunity to challenge the scientific advice directly, apart from through the CMO and CSA. Interestingly, the only time that I remember a senior advisor explicitly challenging the scientific advice through SAGE was when Dominic Cummings attended in March 2020 **{JE/102 - INQ000106152 }**. Perhaps building in an explicit challenge role might improve the quality of advice. Alternatively, organising occasional wider meetings on important topics might have helped ministers and senior civil servants to better understand the limitations of the evidence base and associated advice. The need for these measures probably depends on the individuals involved. We were very fortunate to have the CSA and CMO that we did: both had a fantastic grasp of the detail across an enormous range of fields, an ability to see the wider picture and the capacity to distil complex scientific arguments into simple, clear messages.

16.10. Openness leads to better scientific advice, better decisions, and improved public understanding. The secrecy surrounding SAGE at the outset of the epidemic served no useful purpose and should be avoided in the future.

16.11. The scientific community in the UK played an important role in informing the public. This was greatly aided by the Science Media Centre who worked tirelessly to ensure that experts engaged with the media and public alike **{JE/240 - INQ000212181}**. There were added constraints on SAGE members (and no doubt other committee members) related to confidentiality. This often led to other experts (and some non-experts) filling the press and public's need for information. Speed to publication and openness could help reduce this, allowing committee members (who are presumably most knowledgeable) to speak more freely on relevant topics.

16.12. The public were misled about the role of scientific advisors during the pandemic. The Government did not "follow the science", nor should they have as they had to weigh-up many other ethical, logistical, and economic constraints. On the occasions when the Government clearly did not "follow the science", this led to speculation of rifts between advisors and the Government which made it difficult for advisors to speak publicly as they could be easily dragged into the political sphere. Well publicised clarity on the roles of advisors, policymakers and decision-makers would have been very helpful.

16.13. SAGE and some of its key subcommittees (such as SPI-M and SPI-B) moved into action quite quickly. This basic structure had been planned for and generally worked smoothly (though the role of NERVTAG when SAGE was constituted was less well defined). As the scope of the epidemic widened (as it took hold in the UK) it was clear that there was a need for other subgroups on specific topics (e.g. Social Care). Having many subcommittees in place before the pandemic would have been unnecessary and wasteful. Going forward, it would be important to review which committees are key and would be needed immediately when any event occurs. It is important that these are kept relevant and ticking over. The list of essential advisory committees could be the same basic few that were in place before the pandemic (e.g. SPI-M and NERVTAG), but this is not necessarily the case.

16.14. SAGE and many of its subgroups sat for over two years. The subgroup of SPI-M that deals with medium term projections continues to meet more than a year after all

measures have been lifted. There is an argument that some of the roles and functions of SAGE and its subcommittees should have been handed over to UKHSA and other bodies earlier than it was. Where possible responsibility should be passed onto UKHSA and other relevant bodies as soon as possible, as this is the more sustainable way of working.

16.15. Governments need urgent scientific advice during an emergency. There are many ways to organise this. Some countries relied almost exclusively on their public health system (their equivalent of the UKHSA) leaving much academic expertise outside the formal advice channels. Others, such as the model adopted by SAGE tried to integrate government scientific advisors, the national public health agencies, and academics. Some countries kept the scientific advisors and decision-makers separate (the approach of SAGE and JCVI), whereas others held joint meetings between scientists and politicians. Getting the balance correct is important and different solutions will fit different countries, depending on the strength of academia and public health bodies in different areas and the importance placed on independence from political interference. It is interesting to note that there were even significant differences in the way advice was handled in the different countries of the UK. The Inquiry should certainly look at different models and how they were perceived to have performed. In doing so, however, I think it is important to have some clear criteria as to how to evaluate them, such as what were they trying to achieve, and to contact those who served on these bodies and those who used their advice. Relying on second-hand information from individuals who had little or no direct experience of these bodies and have not done a formal evaluation of them could be very misleading.

International collaboration

16.16. Individual scientists were in contact with colleagues in other countries throughout the pandemic. In addition, the use of pre-prints and social media greatly sped up the sharing of scientific information, though this was, of course, largely not peer-reviewed. The main period when data was lacking and potentially confusing was the initial weeks of January 2020. Ensuring strong national and international public health bodies and that regulations and incentives are in place to share accurate information is critical. It is salient to remember that South Africa were widely praised for their promptness in

alerting the world to the Omicron strain. They were also immediately hit by travel restrictions.

Surveillance issues

16.17. We had poor surveillance at the outset of the epidemic, though this is by no means atypical. Indeed, we were partly relying on existing surveillance systems which miss most respiratory infections as few people with a respiratory infection seek health care and even fewer are tested for the cause of their disease. Thus, the overwhelming majority of cases were missed, allowing the epidemic to grow silently through February and early March. This was compounded by very significant delays from infection to the reporting of the cases that were captured by the system. All this led to very poor situational awareness and almost certainly contributed to a lack of urgency that persisted right up to mid-March 2020. We should build on the later successes of the pandemic (which transformed our ability to track the epidemic) and improve our surveillance system, particularly in the community. Hospital-based surveillance is important, but it is slow, due to delays, patchy, and gives a biased picture of the underlying epidemic. We should put significant investment into developing community surveillance systems, including the use of improved diagnostics. The ONS Coronavirus Infection and REACT studies were transformative in this regard. Taking the best elements of these studies and adapting them for future surveillance would lead to a step-change in our understanding of the burden and spread of endemic diseases, provide a platform for innovative control measures to be evaluated, and ensure that we are better prepared for the next pandemic.

Contact tracing, quarantine and isolation

16.18. Many countries that maintained low (or zero) incidence of COVID-19 during the pandemic had much more stringent quarantine and isolation policies than we designed and implemented. In effect these countries imposed higher costs on high-risk individuals (cases, their contacts, and often the contacts of the contacts, as well as travellers), but in doing so imposed fewer costs on society as a whole. There has been little public discussion about whether we got this balance right.

Access to and availability of data

16.19. There was patchy access to data during the first few months of the pandemic that prevented some groups from contributing fully to SPI-M-O at the time. If external expertise is to be used, then there is a need to rapidly put in place measures for these groups to access appropriate data in a secure way. The Monkeypox epidemic of 2022 highlighted that the gains made during the pandemic were quickly reversed. Despite the urgency of the situation and the potential for a global pandemic, access to data was not forthcoming. There is clearly a need to put in place measures to ensure that appropriate, secure access to data happens quickly for all relevant researchers (e.g. those on SPI-M) along with triggers and procedures for allowing this. SPI-M and UKHSA are evaluating options for data access at the time of writing.

16.20. Some of the major scientific advances during the pandemic came from data-linkage studies, often involving linking data held by different organisations, such as medical records (NHS), testing history (PHE/UKHSA) and mortality data (ONS). Examples include the rapid assessment of the different vaccines **{JE/241 - INQ000212182}**; the quantification of the effect of different factors determining the risk of severe disease **{JE/242 - INQ000212183}**; and the rapid assessment of the relative severity of the different variants **{JE/57 - INQ000212186}**. The UK has significant expertise in this area and great potential, given that we have a unified health system, a large population and strong capabilities in data analytics. These advances should be capitalised on to improve the health of the nation generally. The pandemic showed what was possible with better access to and sharing of data within and between organisations. Unfortunately, these gains have now been reversed with a return to highly cautious and defensive attitudes to data sharing and access.

The use of modelling

16.21. Models are useful tools during an epidemic – indeed, they are the only way to quantify the potential consequences of different policies. There are, however, many issues with the use of models for policymaking. These are not unique to this crisis, and many of these issues arose during the BSE/CJD and Foot and Mouth Disease epidemics, as well as in the use of models for informing climate policy. Projections from models are

inherently uncertain, particularly during the early stages of a crisis, when the demand for these projections is often greatest. Model results need to be treated with caution and their limitations understood by decision-makers, the press and public alike. They should also be handled appropriately. If action is taken to avoid a particular scenario, then comparing subsequent real-life data to that scenario is meaningless.

16.22. Models need to be explicit and reproducible. Along with a comprehensive description of the models, the underlying assumptions, data, and the code, need to be publicly available and properly documented. This requires proper resourcing and inevitably builds in delays. Transferring data analysts, computer scientists and modellers from other fields can help during a crisis and was done via the Royal Society's RAMP initiative. Building surge capacity in crucial areas (including modelling) should be an important feature of pandemic preparedness.

16.23. There are considerable risks in relying on a single model given the uncertainty in any model's projections. Some reassurance can be derived from comparing the results of independent models to each other. This is the basis of SPI-M's approach, and it should be retained for future crises as well as for routine decision-making where models are necessary. There are a few instances where this was not followed. Unusually, JCVI did not follow this approach during the pandemic, despite the availability of different models through SPI-M-O. Furthermore, there was little or no independent comparison of models of transmission in hospitals and care homes. In future pandemics, efforts should be made to ensure that there is adequate modelling support targeted at high-risk settings, in addition to the community.

16.24. Without an explicit model, then decision-makers are left with intuition or ideology to guide their actions. We should, and can, do better than this.

16.25. The lack of explicit economic analysis was an issue. Only one side of the argument – the public health side – was in the public domain. This cannot be conducive to a balanced debate. The same standards of openness and reproducibility should apply to all models being used to inform Government policy.

16.26. There is a need for better integration of epidemiological and economic models. This is not straightforward, however, as the outputs from epidemiological models (cases, hospitalisations, and deaths) are not direct inputs to macroeconomic models, such as 'Social Accounting Matrices', in the case of 'Computable General Equilibrium' models.

Although some attempts have been made to integrate economic and epidemic models, (e.g. **{JE/243 - INQ000212184}**). The complexity of these unified models is likely to limit their use in future pandemic decision-making. Nevertheless, better collaboration and cross-fertilisation between these disciplines would help better prepare the UK for the next epidemic, including helping to determine how much investment should be made in improved preparedness and surveillance.

16.27. Behavioural science is also critical during a crisis. However, the inferences from this field tend to be qualitative in nature. Epidemic models make assumptions about behaviour, for example, contact patterns between individuals in the population and uptake of different interventions. It has not been possible to accurately quantify behavioural predictions for use in models, which has limited their predictive power. Further work integrating epidemic and behavioural models would be helpful. The data collected over this epidemic should serve as a good starting point for such research.

Research

16.28. The UK took a leading role in many of the scientific breakthroughs during the pandemic. This was for many reasons, including a strong biomedical sector, a nationalised health system that allowed large scale studies to be undertaken, a culture of evidence-based-medicine and evidence-based-policy-making, and a receptive Government that was willing to fund large-scale research projects during the crisis. The list of UK scientific achievements is very impressive, ranging from the SIREN study to the ONS Coronavirus Infection Survey; COG-UK to the development of the Oxford/AstraZeneca vaccine; and many more. However, there were significant gaps in the research portfolio – most notably around the non-pharmaceutical interventions and testing, tracing, and isolation strategies. Many interventions were introduced and then removed or altered later in the epidemic, but almost none of these were subject to the rigorous evaluation that is required for pharmaceuticals. This is particularly striking given the UK's strength in large-scale pragmatic clinical trials (as exemplified by the RECOVERY and PANORAMIC trials), the enormous cost of these interventions and the fact that we still need measures to try to avert the future burden of COVID-19. The failure to undertake rigorous trials of NPIs and testing policies during the pandemic was a wasted opportunity. There is still a need to get answers for many of these

questions, however. We should put in place rigorous methods to test the effectiveness and cost-effectiveness of interventions now to try to reduce the ongoing burden of COVID-19 and other respiratory infections (examples might include mask-wearing and testing programmes in hospitals and care homes, but there are many more). We should also ensure that non-pharmaceutical interventions **('NPI's')** and other government interventions are properly evaluated using experimental designs when they are needed next. This lack of high-quality data on NPI's meant that we could never properly answer basic questions on the effectiveness of different measures. We should not have been in that position and should strive to avoid it in the future.

16.29. The Government did make significant amounts of funds available for research during the pandemic. While this was generally successful, there were some issues that should be addressed in the future. The UK Research and Innovation scheme that was set up at the outset of the pandemic had a very quick and streamlined application process. However, funding available through this scheme only lasted for a maximum of 18 months. A grant obtained through this scheme funded much of the LSHTM modelling efforts. However, the epidemic lasted for longer than two years, yet extensions to the scheme were not available. In addition, it was almost impossible to recruit and train new staff during the pandemic. Instead, experienced staff were moved onto COVID response work from other projects. These other projects were then put on hold. Most funders allowed this, but almost exclusively gave one-year no-cost extensions, meaning that staff had to transition back to these projects after 12 months. The net effect was that modelling resources were stretched extremely thin during the second and third years of the pandemic, despite very high demand for modelling services (**e.g.** the "Roadmap out of lockdown", the Delta and then Omicron waves). This put an intolerable burden on the few staff still working on the COVID response. More flexible and longer-term funding arrangements are needed for future epidemics. Indeed, it is salient to note that LSHTM's modelling team also received funding from the EU through its Horizon 2020 Programme. This grant was approved even faster than the UKRI grant but was three years in duration. Hence, much of the LSHTM modelling work being undertaken to help UK decision-making in the second half of the pandemic was funded by the EU. This longer duration of contract also allowed some of the essential wash-up and reflective work to be performed after the acute phase of the pandemic ended. There is a need for much more of this work. Much of the data that was collected during

the pandemic has been under-utilised. There is also a store of experience and knowledge that should be exploited before it is lost and dissipated.

Human resources

16.30. Staff worked under extreme pressure for many months – even years, in some instances. In protracted crises there should be mechanisms (including funding) in place to bring in new staff to ensure continuity of effort. There were significant issues of burn-out with some staff that has had a lasting effect on them and their careers. In addition, the standard metric for academic evaluation is the number and quality of peer-reviewed publications. However, much of the work undertaken did not result in peer-review publications as it was too specific to a particular UK-related problem, quickly became outdated and it was too time consuming to push papers through the peer-review process. Hence, many researchers have little to show for their efforts. Improving incentives and rewards for academic researchers working under such circumstances are necessary. Finally, we are currently in danger of dissipating and losing the cadre of individuals who now have significant experience of outbreak control. We should actively avoid this. Longer term network grants and improved career structures are needed. This should involve academia and funding agencies working more closely with UKHSA and other authorities to help bring this about.

16.31. The extreme pressure that the CMMID group were under meant that it was impossible to take on other work. With a few exceptions, we did not have the capacity to offer significant support to other countries, particularly those from Low- and Middle-Income Countries **('LMICs')**. I am sure that other groups were in a similar situation. Many of the low-income countries were relatively spared, as they have comparatively few elderly and high-risk individuals. The next pandemic may not be focussed on the elderly and high risk. Better provision needs to be in place to ensure that the UK's scientific resources can be employed to improve global health.

Structural issues

16.32. The UK has a large elderly population and an unequal society with high rates of poor health in many parts of the population, including one of the highest levels of obesity in

Europe. It also invests less in health care than many of our neighbours resulting in a health service that is severely stretched most winters. It also has amongst the lowest number of hospital and intensive care (ICU) beds per head of population in Europe in 2020 the UK had an estimated 5.89 ICU beds per 100,000 population compared with 19.04 in France and 47.74 in Germany (see **{JE/244 - INQ000212185}**). These issues inevitably contributed to poorer health outcomes over the course of this pandemic. Addressing the structural drivers of poor and unequal health and wealth will stand us in much better stead for the next pandemic, irrespective of what causes it, and bring about improved health more generally.

The long-term impact of pandemics

16.33. COVID-19 is a new disease that was not present in humans a few years ago. It has become endemic, meaning that we will never be able to rid ourselves of it now. Instead, we will have to treat cases and vaccinate in an attempt to reduce its burden as well as implement other measures to protect those at risk. The health and economic costs of this are enormous. It will be a continual drain on resources. Indeed, since the final easing of restrictions in February 2022, there has been between 5,000 and 20,000 patients hospitalised with COVID-19 at any one time – that is about 5-20% of the available hospital beds. The extreme pressure that the NHS is under is partly due to this acute new demand for services as well as the pent-up demand from deferred treatment during the pandemic. There are many important implications for this. First, that we need better ongoing methods to control COVID-19 from improved vaccines and therapeutics to better diagnostics and measures to prevent infections in high-risk settings. Second, we need improved investment in our health care system. If the last year can be used as a guide, then we need roughly 10% more hospital beds now than we did before the pandemic just to **cope** with the additional continual demand from COVID-19. Finally, we now know that the acute costs of pandemics can be crippling to the economy as a whole. We are now starting to realise that the ongoing post-pandemic health and economic costs will also be enormous. We should significantly increase our investment in measures to improve global health security and ensure that we always act swiftly and promptly to stamp out any future threats.

16.34. It is salient to note that we have not yet been able to stamp out Monkeypox since it emerged in previously uninfected countries in 2022. It continues to spread, albeit at very low levels. Our preparedness and response to threats still has room for improvement despite the experience of the COVID-19 pandemic.

Statement of Truth

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: _____30 August 2023 _____