Witness Name: Graham Medley

Statement No.: 01

Exhibits: GM/01 – GM/271

Dated: 4 September 2023

Ref: M2/SAGE/02/GM

UK COVID-19 INQUIRY

WITNESS STATEMENT OF PROFESSOR GRAHAM MEDLEY

Contents

I, **PROFESSOR GRAHAM MEDLEY OBE**, of the Department of Global Health and Development, London School of Hygiene and Tropical Medicine will say as follows: -

Section 0: Introduction:

1.1. I make this statement pursuant to the Covid-19 Inquiry's Rule 9 request of 10 February 2023.

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

Opening Statement

1.3. I have been researching the transmission dynamics and epidemiology of infectious disease since 1983.

1.4. I moved to LSHTM in 2015, having been at University of Warwick since 1993 and at Imperial College London from 1983. Before that I had a year working on fire safety engineering.

1.5. My aims have always been to understand the patterns in the data, and to provide information and understanding to support decision-making. Necessarily, this means dealing with numbers rather than individuals, but I am aware that the numbers are people. The language and approach used by modellers can appear distanced and without humanity, but that does not reflect sympathy and empathy that I feel. The COVID-19 epidemic in the UK was unusual in that I could see the impact of both the virus and the interventions at first hand. Epidemics are nasty, and everybody carries scars and memories from the epidemic – some much worse than others.

1.6. It was my privilege to co-chair SPI-M-O in order to try and get the best evidence in front of decision-makers in time. It was an extraordinary group of people doing extraordinary things. We were all doing it for the benefit of the decisions, and ultimately the UK population. Just as with my experience on SEAC (2004-2016), the Infected Blood Inquiry (2018-2023) and other advice bodies, my aim is to bring knowledge and understanding from the research community into government.

1.7. It is impossible for memories not to be modified with hindsight. I have tried to give full answers to the questions, and separate what I thought at the time from what I think now.

1.8. I have provided answers to all the questions posed below. Where it appears appropriate, I have provided a section titled "Explainer" which is an attempt to give the Inquiry the relevant background and understanding within which the question and my response sits. I have included references that are publicly available.

1.9. To clarify, I have used "decision-makers" to refer to those people or groups (often, but not always, senior politicians) who decide between different policy choices. I have not used "decision" to refer to a particular choice but to the process of deciding. I refer to "policy-makers" as those people or groups (mostly civil servants) who construct the policy options for decision-makers.

1.10. A critical factor in providing scientific advice was the organisation of UK in four nations in which health is a devolved responsibility. I understand that the Inquiry is separately addressing module 2 issues in these administrations. However, I have provided some answers which cross-over into these sub-Inquiries, and have asked that that this response will be shared with those sub-Inquiries.

Section 1: CMMID

1.11. I was director of the Centre for Mathematical Modelling of Infectious Disease (CMMID) at the London School of Hygiene and Tropical Medicine (LSHTM) from 1 January 2018 for three years, although due to interruptions caused by the epidemic, the new co-directors did not take over until 1 April 2021.

1.12. CMMID is one of the (research) centres within LSHTM. It has a small budget from LSHTM resources (£8000 in 2019-20) which is used principally for facilitating meetings and supporting public engagement. CMMID acts as an interaction / facilitation hub for infectious disease modellers within LSHTM, although with some membership externally. Membership means requesting to be included on the mailing list, and there are no other requirements or expectations. The CMMID monthly meeting was attended by ~50 people pre-pandemic, and the mailing list was ~200 people.

1.13. The director of CMMID is responsible to the director of LSHTM and reports to the Centre Directors Forum. All financial transactions are through the LSHTM systems, and centres cannot hold grants or contracts. The director develops the budget and agrees all expenditure. I ran a management team which met monthly to discuss activities and CMMID initiatives. Ideally each centre has a steering committee (including external membership), but this did not meet during my period as director.

1.14. Members are asked to volunteer to run activities, such as workshops, the regular seminar series, the theme activities etc. One of CMMID important roles is to support early career researchers as they develop modelling skills. The research themes are self-regulating and self-organising groups within CMMID, and open and close as interest waxes and wanes. The themes that have run throughout are the more technical (especially computing) and general (including real-time analysis, elimination/eradication), and those which overlap with other centres (including tuberculosis). The full range of activities is available on the centre website [GM/1 - INQ000213026].

1.15. CMMID had no formal interaction with other modelling groups involved in SPI-M-O, although interactions were facilitated through the seminar series and themed workshops. Additionally, academic early careers researchers are not permanent appointments so that there is a constant flow through the centre as new students and postdoctoral researchers arrive each year and typically move on after 3 or 4 years. Dr Eggo (who was on the management team and is now co-director of CMMID) started *IDDconf* in 2017 which is an annual meeting with majority attendance by CMMID but with about 50% attendance from other modelling groups. The great majority of established researchers in infectious disease modelling in the UK are directly connected through collaboration and training, so that most groups have people in them who have worked and trained at other groups. The infectious disease modelling community in the UK is relatively small (several hundred people), so that we personally know a lot of those involved. This meant that SPI-M-O could quickly form as a functioning team.

1.16. During my final year as director of CMMID I relied very heavily on the management team to organise everything. I could see on the CMMID 'slack channels' that the centre was functioning and adjusting to working from home. 'Slack' is a proprietary software that provides a virtual working environment and is and was heavily used by CMMID [GM/2 - INQ000213141] https://slack.com/intl/en-gb/. My main

concern was that early career researchers were able to continue to interact with their peers. I tried to attend as many on-line meetings as I could, although SPI-M-O and SAGE commitments prevented much involvement. A lot of the activities were paused as many members of CMMID transitioned to working on the pandemic and learnt how to work from home.

1.17. I took the decision very early in February 2020 that if SPI-M-O was to function as a group of peers reaching a consensus, then I should, as chair, step away from the modelling work in LSHTM in order to preserve my independence and not be seen to be favouring particular groups. Consequently, my direct involvement in CMMID and its COVID-19 modelling was very limited and I am unable to comment in any detail. I attended a small number of the weekly COVID co-ordination meetings, but as much for personal interaction as providing general comments about SPI-M-O and the general situation as I saw it and learning similar from the other attendees. One of my post-graduate students joined the COVID-19 working group, so I knew something of the way it was working. All papers that were released by CMMID were internally reviewed, and I was a member of the CMMID COVID-19 Working Group in order to read and comment on research within CMMID as it developed from the viewpoint of research process and presentation rather than policy development. For example, I was included as explicit co-author on one paper because, during review, I suggested that it would be better to present additional/relative value of backwards contact tracing rather than the simple value [GM/3 - INQ000213249].

Section 2: CMMID COVID-19 Working Group

2.1. My understanding is that Dr Eggo set up the CMMID COVID-19 working group with input from Dr Kucharski, Dr Funk and Professor Edmunds (all of whom are known to the Inquiry through SPI-M-O membership). Dr Eggo had been heavily involved in the Ebola response in 2014-16 and had experience of what would be required and the speed at which it would have to operate. Professor Edmunds was closely involved in some of the research activities, but less closely involved in developing the structure and *modus operandi*. Dr Eggo was keen to ensure that the researchers who joined were supported and that their work would be recognised, hence the formation of the CMMID COVID-19 Working Group to confer joint scientific publication authorship for all involved.

2.2. I was not closely involved in CMMID COVID-19 work and am not able to provide full information that the Inquiry is requesting.

2.3. I can only comment on the value of CMMID in relation to the other modelling on SPI-M-O, rather than the internal workings of the group; because of my role as chair of SPI-M-O I kept some distance from the details of the internal activity due to both lack of time and being independent of individual groups. From my perspective all the work bought to SPI-M-O was given equal weight and the value and contribution of any piece of work was decided during discussion on SPI-M-O. Because of the number of people at CMMID, the number of papers and amount of work presented at SPI-M-O was more than most other groups. The contribution of modelling to policy development (including that from CMMID) was mostly through SPI-M-O via consensus statements and SAGE papers.

2.4. Early in the epidemic I believe that some modelling was being directly communicated with government, CMO and GCSA rather than through SPI-M-O, although I am not aware of what was communicated. I believe that Professors Edmunds (LSHTM) and Ferguson (Imperial), who were on SAGE in their own right, were contributing to such discussions during February and March 2020, although I think that this lessened after the initial period. Members of SPI-M-O were explicitly asked to keep the SPI-M-O secretariat informed of all interactions with government.

2.5. Note that local collaborations based on research rather than policy might have informed local decision-making within healthcare. For example, I am aware that Dr House (Manchester) was collaborating with the Manchester Royal Infirmary and that Dr Eggo was collaborating with community leaders in north London.

2.6. The Inquiry will have to get details of the internal workings of the CMMID COVID-19 Working Group from others, principally Dr Eggo and Professor Edmunds.

2.7. I was not involved in creating any of the tools that CMMID COVID Working group produced. The Inquiry should ask others for accurate, first-hand information. I suggest starting with Dr Eggo and Professor Edmunds who will be able to direct the Inquiry appropriately.

2.8. I was not directly involved in CoMix; I believe that Professor Edmunds was the principal investigator. He will be able to provide the details requested.

Explainer: Contact patterns

2.9. I can explain the importance to modelling of having the CoMix and other such data on behaviour. In the simplest models, the assumption is that everybody has the same experience as everybody else. In particular, if I meet 10 people tomorrow, then those people are drawn at random. If the prevalence of infection is 10%, then on average one of the 10 people I meet will be infected. In reality, of course, we do not meet people at random. We tend to meet the same people repeatedly (friends, family, household), and when we do meet people outside of our immediate network, they tend to be similar in certain respects (for example, age and occupation). The most notable dimension on which our contacts cluster is age. The *polymod* study published in 2008 was a big step forward in making models more realistic as provided consistent, large, random data sets on contacts by age [GM/4 – INQ000213260]. Since 2008, these data have been used repeatedly, but have not been updated. CoMix was set up to provide more timely data, and to monitor the impact of the epidemic on behaviour. The study successfully recorded the changes in average contact rates over the period of the epidemic and provided international comparisons.

2.10. It has long been recognised, especially for sexually transmitted infections, that as well as knowing the average rates of contact, it is important to know the contact rates of the contacts, as this creates networks of contact. For respiratory transmitted infections (or at least where transmission does not require the same degree of intimacy as sexual transmission) the average contact rate by age has proven sufficient to capture the transmission dynamics, at least for high-level policy purposes. However, there are indications that during the epidemic, network dynamics became more important, with people restricting contacts to a smaller set of repeated contacts. This is intuitive, as for example, working from home and LD regulations meant that random contact during travel was greatly reduced. CoMix was not designed to capture these networks.

2.11. During epidemics, normal human behaviour changes, either as a result of government intervention or spontaneously. There was no data on how behaviour in the UK responds spontaneously to epidemics, nor on how behaviour responds to government intervention when the interventions have not been used before. CoMix was essential to be able to get insight into what people actually did during epidemics.

2.12. There is clearly a need to be able to understand contact patterns better, and especially during the next epidemic. This needs considerably more research to develop. The NHS app developed for tracing purposes (i.e., recording contacts) clearly has the necessary technology, although considerable barriers remain to be addressed. ***[explainer ends]***

Section 3: SAGE and sub-group interactions

3.1. The Inquiry's understanding of my role and participation is correct – I was academic co-chair of SPI-M from October 2017, co-chair of SPI-M-O from January 2020 until the last meeting in February 2022, and have since returned to co-chair SPI-M. SPI-M did not meet whilst SPI-M-O was meeting. I resigned from co-chair of SPI-M in June 2023.

3.2. SPI-M-O is a sub-group of SAGE tasked with addressing the scientific evidence that comes from considering the transmission dynamics of infectious disease, in this case SARS-CoV-2. Mathematical modelling of infectious disease is a specialist subject, so the formation of a sub-group focussing on this area is sensible and required. See explainer below for more background on the subject and at paragraph 3.42 for more details on my role. SAGE and its sub-groups relied on their secretariats (groups of civil servants) to provide infrastructure and to co-ordinate between sub-groups.

3.3. The sub-groups provided the bulk of evidence reported to SAGE. Independent attendees bought interpretations of external evidence (e.g., from the scientific literature) and their experience, but most of the written, UK-specific material was generated by sub-groups. As a consequence, the SAGE consensus was largely a composite of the interdisciplinary discussion of evidence generated by the disciplinary expertise on sub-groups. The sub-groups' consensus statements were provided to SAGE prior to meetings. The role of chairs was to primarily answer questions and ensure that attendees, and particularly CMO and GCSA understood the contents.

3.4. I can only describe the relationship of SPI-M-O with other sub-groups from my standpoint. Each sub-group had its own disciplinary expertise. Having good lines of communication between each discipline is essential to developing

interdisciplinary evidence. My understanding is that of the sub-groups listed on the SAGE website, only NERVTAG and SPI-B were meeting prior to 2020. Cross-membership of Professors Edmunds and Ferguson between NERVTAG and SPI-M-O facilitated communication, whereas the relationship between SPI-B and SPI-M-O was non-existent initially but developed through to summer 2020. As other sub-groups were set up, I tried to ensure cross-membership usually by nominating a member of SPI-M-O as the representative, and often attending the meetings myself as well. The SPI-M-O secretariat were also attending some of these other sub-group meetings. Below I describe the relationship with each sub-group in turn. There is more detail for some of the sub-groups in paragraph 3.97. Note that some of the groups were known by multiple names. I have given my own classification into those sub-groups focused on specific fields and those sub-groups focused on particular issues which required multiple disciplines to address. I have dropped academic titles in the descriptions. There were also research projects and collaborations on-going between the members of different sub-groups. I am not aware of all of these and have only described the 'official' interactions, but these academic, research interactions were important.

Disciplinary sub-groups

i. NERVTAG. There was a good relationship because of cross-members Edmunds and Ferguson. As co-chair of SPI-M I had been invited to NERVTAG meetings prior to 2020.

ii. SPI-B. Unfortunately, there was no interaction between SPI-M and SPI-B prior to 2020, and the members (and area of expertise) were new to me in February 2020. The relationship developed rapidly. Brooks-Pollock attended several SPI-B meetings, and SPI-B members attended some SPI-M-O meetings: Yardley four and Rogers two, as recorded in the minutes. However, it was co-working in the interdisciplinary sub-groups and that provided the most positive and in-depth communication.

iii. PHE serology working group. Baguelin and Van Leeuwen attended, although there was sufficient academic linkage prior to the epidemic that significant developments went through existing networks.

iv. CO-CIN. There was no cross-membership. The CO-CIN reports submitted to SAGE were very useful and the CO-CIN group made their data available.

Personally, I relied on SPI-M-O member Read for insight into the data. I also had conversations with CO-CIN lead Semple when there were issues that needed to be discussed, in particular the number of nosocomial infections.

v. Environment Modelling Group (EMG) – SPI-M-O member Hall attended some of these meetings and there was communication within the inter-disciplinary sub-groups. On reflection, the communication between EMG and SPI-M-O could have been more regular and more complete, although I am completely confident that had special need arose, I and SPI-M-O would have been alerted.

Inter-disciplinary sub-groups

vi. Children's Task and Finish Group (Children's TFG) ('SPI-Kids') – this group was initially organised by the SPI-M-O secretariat and there was considerable cross-membership especially with SPI-B. School closures were one of the first issues that SPI-M-O was asked to consider as an intervention, for example, see email 13 February 2020 [GM/5 - INQ000213271]. I attended early meetings but withdrew as SPI-M-O was well represented.

vii. Hospital Onset COVID-19 working group (HOCI) (Nosocomial transmission sub-group) – This sub-group was initiated by SAGE 21 31 March 2020 largely as a result of the combined CO-CIN / SPI-M-O analysis detailed in paragraph 3.97 [GM/6 - INQ000213282]. I introduced Robotham to the SPI-M-O secretariat 15 April 2020 and she developed a modelling sub-group (including Dr Knight from LSHTM and Professor Cooper from Oxford) [GM/7 - INQ000213295]. I initially attended the meetings but withdrew as Robotham became more engaged.

viii. Ethnicity – SPI-M-O member Eggo attended these meetings.

ix. Social Care Working Group (Care homes sub-group) – This sub-group was initiated during the first wave when the impact on care homes (CH) was manifest. SPI-M-O member Hall was closely involved and acted as co-chair. SPI-M-O members Danon and Finnie also attended. I initially attended meetings but stepped back as Hall became more engaged. I was initially concerned that the detailed modelling required for LTCF was sufficiently linked into SPI-M-O work, as in email 25 April 2020 [GM/8 - INQ000213307] [GM/9 - INQ000213319]. The PHE Modelling Cell (including SPI-M-O members Hall, de Angelis and colleagues) took on the bulk of the modelling work in government facilities (including social care

and prisons), and there was a risk it would dissociate from SPI-M-O, but this was avoided.

x. It became clear during April 2020 that there had been little planning for an epidemic in LTCF. The data were initially very limited, and personnel seem to be drafted in from other branches of government; see this email thread 3 May 2020 [GM/10 - INQ000213027]. During May 2020 I became increasingly concerned that Hall should have been co-ordinating data and modelling work, and not chairing a sub-group of SAGE. This is not to say that I believed that he had done a bad job (quite the opposite), but that it should be co-chaired by a senior civil servant, as discussed with FCO CSA in email thread 22 May 2020 [GM/11 - INQ000213038]. However, Hall was prepared to continue, and did not ask me to push harder to get a replacement which I did offer, and I believe he continued to chair/co-chair the group until at least April 2021.

xi. Enduring prevalence sub-group – The SPI-M-O sub-group on spatial patterns (see paragraph 3.14) was chaired by SPI-M-O member Gog, and she and I also attended the Enduring Prevalence sub-group.

xii. Higher and further education sub-group – This was attended by multiple SPI-M-O members. I played a relatively minor role.

xiii. Mass testing sub-group – I attended and contributed to the writing of the report which went to SAGE 53 27 August 2020 [GM/12 - INQ000120552] [GM/12A – **INQ000229533**].

xiv. Vaccine sub-group – I attended along with SPI-M-O members Gog and Ferguson.

Explainer: Infectious disease transmission dynamics and models

3.5. Infectious diseases show several differences from other health threats in terms of their propagation and natural patterns. There are three main features of infectious disease dynamics (IDD). First, the risks are dynamic, i.e., change with time, because incidence (the rate at which people become infected) is a function of prevalence (the number of people who are infected). The number of people infected last week largely determines the number infected this week, so that if measures were put in place to prevent infection last week, there will be fewer infections this week. Both of these effects were evident during the epidemic with

rapidly increasing waves, i.e., more infections last week, resulting in more infections this week, resulting in more infections next week, and the impact of NPI which slowed the rates of infection.

3.6. Second, infectious agents evolve in response to both the biological processes and the interventions put in place. There were three significant variants in the UK (alpha, delta and omicron) which changed the IDD.

3.7. Third, infection and disease are different, but it is infection that must be considered as it is infections that drive the IDD. A large proportion of infections are asymptomatic, and are normally unobserved. Additionally, the delays between infection, symptoms, hospitalisation and death mean that the epidemic is never actually observed in real time – the data are always lagged. The epidemic was the best observed epidemic in that the ONS CIS, REACT and the availability of testing meant that we were closer to real time observation, but nonetheless the lags in the data were significant. We can only confidently say what has happened in the past, not what is happening today hence the need for 'nowcasting'.

3.8. The consequence of these three features is that mathematical models of transmission dynamics are the natural tool. Whilst models are essentially mathematical, they have become increasingly operationalised within computer code. Modelling IDD is essentially interdisciplinary, requiring knowledge of (at least) mathematics, statistics, biology, clinical sciences and computer science. ***[explainer ends]***

3.9. My understanding is that the role of SAGE and its sub-groups is primarily to provide CMO/GCSA with the scientific evidence which they then form into advice and take into government and decision-makers. SAGE and its sub-groups developed a consensus in the sense that what was written was agreed by all not to be wrong. Variation in opinion or evidence was included either explicitly or as a reduction in strength of evidence and confidence in a statement, both during discussion and in the consensus statements. There was, for example, never any vote taken to decide the consensus view. The delays created by requiring a consensus were at most two calendar days. SAGE and SPI-M-O meetings were called and held the same day if required, especially in the early stages and before the 'battle rhythm' became established.

3.10. It should be noted that the work that SPI-M-O were doing was not research, and members had to be reminded that to focus on what is already known and put it into context for policy. Data, information and understanding accrued rapidly during the epidemic, with members being significant drivers of that research effort, and SPI-M-O was very alert to the changing knowledge-base. Application of the knowledge-base to answering policy questions is a different activity from research in that members were not generating or testing hypotheses, but synthesising what is already known.

3.11. The delays of the consensus process do not prevent advice or opinion being given at any time. On several occasions I gave my opinion on a question, and then offered to take it to SPI-M-O for a consensus view. There is always a trade-off between the immediacy and robustness of evidence and advice. Decision-makers must decide this trade-off. A consensus view always carries more weight than an individual opinion, but takes more time to generate.

3.12. I cannot comment on the impact of the delays of the consensus process on the likelihood of advice being considered. However, I was told several times that the consensus process generated evidence that carried much weight. Especially early in the epidemic (until end April 2020) I felt frustrated by some of the delays in response to the immediate threat at the time, but on reflection think that the consensus process produces better evidence.

Specific issues raised by the Inquiry:

3.13. I cannot comment on the relationship with senior politicians. I personally did not meet or speak with any politicians, and to my knowledge none attended SPI-M-O or SAGE meetings. All the meetings had a virtual / dial-in capability, and after 23 March 2020 all meetings were attended remotely. Whilst the secretariats took great care to ensure that the attendees at the meetings were invited, it is not possible to know who is at the end of, for example, a telephone or a zoom call. I was called to the House of Commons Science and Technology select committee twice and gave oral evidence, and on the second occasion provided written clarification. I was aware that CMO and GCSA briefed MPs as well as meeting with ministers. I do not think that it would have been appropriate for me to speak directly with politicians, partly to avoid confused interpretations of the evidence, and partly because modelling should not have a primacy in terms of the evidence base.

3.14. The meetings I attended as part of SAGE were where I was invited by government to attend. I did not take minutes or record attendance. The volume of work that SPI-M-O was asked to undertake necessitated separate sub-groups of SPI-M-O. These meetings were organised by the SPI-M-O secretariat who also attended. Aside from the main SPI-M-O meeting on Wednesdays, there were two weekly sub-groups that I regularly attended on Tuesdays. The first of these was the 'space group' which Professor Gog chaired, and at which different groups presented their work on the current state of the epidemic and its spatial patterning. The second, which I chaired, was about the estimates of the reproduction numbers, growth rate and MTP. The aim of both meetings was for members to have a technical discussion about the robustness of their results and to come to a consensus view which was then taken to the main SPI-M-O meeting. Throughout the epidemic there were many *ad hoc* sub-group meetings to discuss issues such as data and RWCS, and these all reported back to the main meeting for overarching discussion.

3.15. The CMMID COVID-19 working group did not have primacy on SPI-M-O. In the early stages of the epidemic, there was a less equal balance of evidence from different groups on SPI-M-O, but this quickly changed after April 2020 when the data were available to all groups (see Section 7).

3.16. As far as I am aware, SAGE and its sub-groups considered all international data that was available to members. This was facilitated by personal relationships that by-passed normal publication processes. Evidence from international sources must be weighted by knowledge of differences between countries. For example, I recall that it was reported to SAGE in February 2020 that oxygen provision for hospitalised cases was very high in Singapore, but this must be weighted by the normal healthcare practice of providing oxygen to patients and apparently the clinical threshold for oxygen provision is much lower in Singapore than UK.

3.17. One of the principal functions of SAGE is to inform CMO/GCSA. CMO in particular has access to other evidence advice, not least directly from PHE/UKHSA. SAGE is not the sole source of evidence available to senior advisors and officials, although SAGE does have the strength of being able to create an interdisciplinary consensus that is difficult to achieve otherwise. I do not have a view on whether the balance of evidence coming from SAGE versus other sources was correct, given that it was not run for my benefit. There was certainly no shortage of clinical input into SAGE given CO-CIN and the clinical training of many attendees, and given that SAGE was

not informing clinical decisions but Government policy. SAGE was asked whether additional disciplines should be represented, and the attendees changed during the epidemic. The public health directors were represented on SAGE. Personally, at no point did I feel that, for a scientific evidence and advisory committee, that there was a paucity of evidence from any discipline.

3.18. I had some influence in the membership of SPI-M-O, and gave diversity of background some consideration, although my overriding aim was to ensure that there was diversity in expertise and experience rather than where they trained. All members of SPI-M-O had substantive academic posts related to the transmission dynamics of infectious disease. I was more concerned with diversity of technical expertise than any other dimension.

3.19. When suggesting members for SPI-M-O, I purposefully looked to ensure that the balance of experience would provide challenge. I was mindful throughout the epidemic of having to avoid both the committee becoming too insular and mutually supportive (and in danger of 'group-think'), and it becoming too competitive and challenging to reach a consensus. Academic research is highly competitive and there was plenty of challenge of each other's work and results on SPI-M-O, although it was rarely destructive criticism. Given that all the work SPI-M-O members presented was being put into the public domain, there was a lot of challenge from outside of SPI-M-O. In terms of the modelling evidence, I am content that there was sufficient challenge.

3.20. SAGE was different in that the main purpose was to bring together the work of the sub-groups in front of CMO/GCSA. There were attendees who were appointed to comment on the SAGE meetings. The SAGE papers going into the public domain was also an opportunity for external challenge.

3.21. It is likely that, for individual issues, both SPI-M-O and SAGE could have produced different evidence and advice, but I do not think that the process lacked sufficient challenge.

3.22. I was personally clear about my role chairing SPI-M-O. I had been academic chair SPI-M for almost three years, and from that experience understood that my role was to facilitate the group of modellers to produce a consensus of evidence for policy development. I had developed a good working relationship with the SPI-M secretariat, which was critical from January 2020 onwards. I was less clear initially

about my role on SAGE, although discussion with the secretariat, and the interactions at the meetings made it clear what was required. The details of my role are further addressed in paragraph 3.44.

3.23. I was influenced by the Hine review of the 2009 swine 'flu pandemic' which makes a clear distinction between evidence and advice (recommendation 14 and para. 4.60) [GM/13 - INQ000213060]. I made it clear to the secretariats that I regarded my role was the synthesis of disciplinary-specific evidence on SPI-M-O, and that SAGE was where evidence across disciplines was bought together in front of advisors. The primary role of SAGE was to ensure that the advisors (including CMO and GCSA) understood the breadth of evidence, and, critically, the uncertainty associated with it. Nobody disabused me of that view. See also paragraphs 5.7 and 5.8.

3.24. I believe that the chairs of the other subgroups on SAGE had similar roles, based on conversations and my observations. The roles of the CMO and GCSA in regard to SAGE were clear from the conduct of the meetings. One of the principal functions of SAGE was to ensure that CMO and GCSA had the information they required, so to a large extent they determine the constitution and role of attendees to suit that purpose. There were a large number of civil servants from across government who listened into both SPI-M-O and SAGE meetings as observers, but I got very little feedback as to whether the meetings were useful or informative for them. The exception was SPI-M-O meetings from January 2021 at which the Cabinet Office attendees gave guidance and feedback. Personally, I was very happy to see that government advisors were listening in. I would have been very happy to answer any questions that might have arisen. I used the distinction between evidence and advice on SPI-M-O meetings on several occasions in order to help members understand that our role was to provide evidence, as free as possible from advice and advocacy.

3.25. The clear distinction between evidence and advice was challenged by a number of situations, not least of which when the evidence is sufficiently compelling that the advice creates itself. The distinction between evidence and advice can become blurred in the political sphere, where the strategic presentation of evidence plays a crucial role in constructing persuasive arguments. There were cases where the evidence that SPI-M-O generated was misrepresented and misinterpreted as 'spin'

in order to justify a preferred policy choice. Regrettably, there were also cases where the process of creation of evidence was misrepresented.

3.26. The question of how much weight to put onto different forms of evidence and advice, with appropriate consideration of uncertainty, is critical. Given my role, I cannot have an opinion on the weighting of different forms of evidence that ministers considered.

3.27. There was a confusion throughout the epidemic about the concept of uncertainty, which is different from a lack of knowledge or understanding. My conjecture is that policymakers and decision-makers confused statements in consensus statements about uncertainty as scientific advisors lacking understanding or knowledge which might have meant that the early documents, which are necessarily uncertain, were dismissed. I say more about the role of uncertainty in paragraphs 11.12 to 11.24. In January-March 2020 I do not think enough weight was put on SAGE evidence. It is necessarily true that at the start of an epidemic of a novel pathogen, that evidence will be largely based on interpretation of scant data by experienced and expert judgement, and given the unavoidable lack of data, modelling evidence carries more weight. SAGE is set up to provide a balance consensus opinion of the evidence and considerable uncertainty. Given that an epidemic is essentially a natural disaster, there is an inevitable reality, which has to be weighed against other considerations, including political issues, by decision-makers. My view is that the reality of the epidemic, via the SAGE process, was not given sufficient weight initially,

3.28. There was a sense that government strategy was being created 'on the hoof' during February and March 2020. I was fully expecting to be answering questions from across government about the SPI-M pre-pandemic preparedness document [GM/14 - INQ000212124], but that was never mentioned or queried. The role of that document is to provide the basis for strategic decisions in the early stages of an epidemic, when there is inevitably scant firm evidence and much uncertainty. The document makes no mention of generalised NPI ("lockdown") to prevent an epidemic. The strategy pre-2020 was to have the epidemic and use NPI to modify the peaks and troughs rather than to stop the epidemic completely.

3.29. Given China's response, which in January 2020 was to close their economy rather than face the epidemic and its consequences, it was likely that generalised NPI would be a widely used intervention in the pandemic. Italy's response in February

2020 confirmed to me that the political and strategic response to the developing epidemic in the UK would include generalised NPI. The word "lockdown" does not appear on my computer until end of February 2020, and its specific definition is still lacking.

3.30. I gained the impression that governments globally were reacting to a situation that was largely unforeseen and that much of the political decision-making was, in essence, gaming between nations, i.e., governments were making decisions based on the actions of other nations rather than independently. This is not surprising given that all countries faced the pandemic threat. There was, for example, considerable public and political interest in Sweden which was characterised as having adopted a different strategic response from the rest of Europe. Figure 1 in this scientific paper suggests that the decision to close schools, for example, was 'global' rather than tailored to local circumstances [GM/15 - INQ000213082].

3.31. From April 2020 onwards the situation changed, and more weight was put on to SAGE evidence. As discussed more in paragraphs 8.43 onwards my view is that the weighting of different forms of evidence was good from January 2021 onwards, but otherwise I cannot say what evidence was used to make decisions.

3.32. I agree with Professor Edmunds' statements that during the first months of the epidemic obtaining resources for modelling to be able to provide evidence was a major distraction. However, I can only give a view from the perspective of SPI-M-O rather than individual groups. At the start of the epidemic, I was focused on ensuring that the modelling teams on SPI-M-O had sufficient data, computational power, and personnel. The data issues are addressed in Section 7. Both computational power and personnel require financial resources.

3.33. Access to sufficient computational power was a critical issue. Whilst doing research, waiting overnight for code to run is not a significant problem, but during the epidemic and given the volume of data that was eventually provided, models required more computing. There was some central (SPI-M-O secretariat level) involvement, for example, making members aware of Dstl, DoD, ECDC and Meteorological Office offers of resources, but groups were responsible for resolving this question themselves. Some groups I know worked with commercial cloud-compute resources, but I do not know how this was financed. Obtaining this resource will have required some investment to get code to run and ensure data security amongst

other hurdles, all of which will have been "a major distraction". I did not keep a record of what individual groups were doing, and, to the best of my knowledge, neither did the SPI-M-O secretariat.

3.34. Ensuring that SPI-M-O had the capacity to provide modelling evidence was a central concern of mine from early on (see email 28 January 2020 [GM/16 - INQ000213093]). This was largely about personnel – having sufficient numbers of well-trained staff who can support the development of models (including testing and documenting code and preparation for publication) and its interaction with data. DHSC had two current grants to SPI-M-O member groups that could be repurposed to support COVID-19 work. LSHTM and Imperial College both had current funding that supported emergency response and additionally had relatively large cohorts of research students that could be redirected to COVID-19. Other groups were smaller and would need to have support. I was invited to join the 2019-nCoV Rapid Response Call Expert Panel on 7 February 2020 and agreed to try and ensure that modelling was appropriately represented, even though I did not really have sufficient time to spare [GM/17 - INQ000213105] [GM/18 - INQ000213106].

3.35. There is a fundamental problem with asking research funders to support modelling feeding directly into policy questions. SPI-M-O was not doing hypothesis-based research, but providing technological answers to operational questions. Funders (including but not only UKRI and Wellcome Trust) were very generous allowing members to re-purpose current grant funding to support government needs, and some grant applications from members were successful. Although this gap between research and operational support was and is unstated, it was clear to me that UKRI did not wish to spend its limited resources on activity which would not hugely advance scientific understanding, but which would provide government with critical evidence; their obvious conclusion being that government should pay for it. On the other hand, DHSC and GO-Science hugely valued the independence of the academics involved in SPI-M-O and appeared reluctant to support research groups directly. Additionally, the process of applying for research grants and various contractual and reporting obligations if they are awarded, represent considerable overheads for researchers who are, ultimately, volunteering to support government.

3.36. Government was aware that more resources were required, and turned to other organisations to provide it, as shown in this email 28 March 2020 [GM/19 - INQ000213130]. Organisations such as the Royal Society and the Royal Statistical

Society were also very keen to be involved and support SPI-M-O [GM/20 - INQ000213142] GM/21 - INQ000213156] [GM/22 - INQ000213157] [GM/23 - INQ000213158] . However, I faced at least three problems with how to respond.

- First, I was co-chairing a DHSC/SAGE committee at their behest, and there was no structure for feeding in work from different bodies, and I did not have the authority to set up the structure. I was keen that all infectious disease dynamics evidence should come through SPI-M-O to add to the consensus (see 28 March 2020 email [GM/22 - INQ000213157] [GM/23 - INQ000213158] [GM/24 - INQ000213194]) but this required that the existing members devote some time to ensuring that the work done was fit for purpose, i.e. involved sufficient infectious disease knowledge.

- Second, the modelling expertise required data, which I did not own and could not give permission to access. I do not know if any academics not attending SPI-M-O gained access to the relevant data.

- Third, I did not know the time frame over which SPI-M-O would be working but did understand that the then members of SPI-M-O would contribute because it related directly to their area of research. I could not rely on modellers and statisticians who were volunteering in an emergency and who might return to their normal work within a few months. The announcement of JBC in June 2020 further complicated the picture because it suggested that government would have this as its primary source of evidence. Had I known that SPI-M-O would be highly active for two years then I would have seen more benefit for getting others more closely involved on a sustainable basis.

3.37. In the event, I believe that the greatest benefit was to individual groups rather than SPI-M-O or SAGE. Several groups boosted their capacity by co-opting external volunteers with specific skills, and my general approach to non-SPI-M-O members offering help was to find a group within SPI-M-O who they could work with. There needs to be a 'standing capacity' of trained people and models to respond in an emergency and trying to fit new groups into the existing structure and provide required support was a distraction. This is not to say that I feel anything but gratitude to those who offered support, and some regret that I did not make better use of them. Hopefully UKHSA will act as the structure in future epidemics.

3.38. The SPI-M-O co-chairs and secretariat understood the individuals who attended SPI-M-O were supported by a large number of scientists, the majority being research students or post-doctoral researchers who diverted from their usual activities. Given the duration of SPI-M-O activity and intensity of some periods, this will have impacted on their careers. In order to give some recognition, we developed the *SPI-M-O Award for Modelling and Data Support* (SAMDS) which was given to over 100 people who made substantive contributions to the modelling effort. I am well aware that this is scant reward for the effort and commitment but is an acknowledgement of the very significant contribution.

3.39. Academics involved in SPI-M-O were not doing research or academic work, but using the results of their academic research to provide support to government. There is some benefit to them and their employers in that they are able to clearly demonstrate that their research is useful and has impact, but this does not recompense the financial cost. Some money was provided to employers to account for academic work not done (e.g., lecturing and marking), but it did not cover the full financial cost. Members relied on their academic colleagues to take on extra duties, and their employers to allow them to be 'seconded' to SPI-M-O. In March 2021 I was asked to continue to co-chair SPI-M-O for another year but indicated that my employer would have to be directly compensated for my time away.

3.40. As academic co-chair of SPI-M-O I was trying to ensure that the modelling available covered what was needed. I did not know until March 2021 how long SPI-M-O would be active, so was unsure how much to ask of members in terms of longer-term modelling development. In the early stages I was cajoling members to work on areas that I thought would become important, but it would have made that task much easier, and I think that the coverage of models would have been improved if I had had access to funding to support different activities. Given the scale of the financial impact of the epidemic, I am surprised that there was not more central support for the modelling.

3.41. One of the consequences of the lack of resources was that SPI-M-O members were distracted from providing the immediate evidence. They largely responded by extending the working day but potentially with a consequent loss in performance and personal toll.

3.42. I was approached to chair SPI-M in 2017 by Peter Grove – a DHSC official who had set up and chaired SPI-M since it's institution. I started in October 2017, and he retired in 2018 [see GM/25 - INQ000213205, GM/26 - INQ000213223, GM/27 - INQ000213236, GM/28 - INQ000213247] [GM/29 - INQ000213248]. My understanding at the time was that being "peacetime" chair of SPI-M implied having a role in "wartime" chairing. In this capacity, I was put on the SAGE experts list 2018. This change also saw the institution of the co-chair system in which the 'academic' co-chair operated outside of government and the 'policy' co-chair operated inside of government. This proved to be very important during the epidemic (see paragraphs 3.47 to 3.68).

3.43. The SPI-M committee was emailed about the "novel coronavirus" 21 January 2020. I had a call with the SPI-M secretariat 24 January 2020 and SPI-M met 27 January 2020 to discuss what modelling had already been done and preparedness for an epidemic if it occurred. SPI-M-O had its first meeting 3 February 2020, following the 2nd SAGE meeting 28 January 2020 [GM/29 - INQ000213248] [GM/30 - INQ000057492], which I did not attend. I cannot recall, and do not have any record, of at which point I was asked to chair SPI-M-O, but I was in place for the first meeting, as I was expecting to be from 2018. The SPI-M-O terms of reference agreed at the first meeting, lay out the initial framework for the committee's work.

3.44. My role was largely defined in the documents referenced earlier in this question, although it did evolve as the epidemic went on. As academic co-chair of SPI-M-O my role was to ensure that the evidence generated was the best available and was understood by SAGE attendees. My principal duties were:

- advise on the membership of SPI-M-O;
- liaise with the SPI-M-O secretariat on the agenda and work required, on the basis of the questions the committee was being asked and the best way to get committee involvement; I was also able to raise other areas that I thought modelling could help address and which might be useful (but which had not been asked);
- chair the main SPI-M-O meetings so that the secretariat were able to gather the information they required and comment on and agree minutes;

- liaise with the SPI-M-O secretariat to finalise the consensus statement, paying particular attention to the uncertainty statements;
- present the consensus statement to SAGE, ensuring that it had been understood;
- attend relevant SPI-M-O sub-group meetings and monitoring performance of sub-groups;
- liaise with the SPI-M-O members and attendees to keep them working as a functional team;
- liaise with others as required to support the members of SPI-M-O to get the information, data etc that they required to do their work;
- liaise with government as required to support their understanding of SPI-M-O and its output;
- join other sub-groups of SAGE and input into evidence documents;
- liaise with public (including media) to explain the work of SPI-M-O.

3.45. On SAGE meetings, I tried to make it clear when I was offering personal opinion rather than SPI-M-O consensus. I was aware that if SPI-M-O was stood down, then I would likely not be attending SAGE, i.e. I was there to represent SPI-M-O consensus.

3.46. Overall, I attended approximately 1700 meetings in my diary 1 January 2020 to 31 March 2022, of which 800 approximately were related to my roles on SPI-M-O and SAGE. I did not organise any of the meetings relating to my roles on SPI-M-O and SAGE, but attended at the invitation of different government bodies. I kept personal notes during the meetings as an *aide-memoire* for actions required of me and others, and as these actions were completed, I discarded the notes. I did not keep a record of the other attendees or how others were recording the meeting. The purpose of the meeting was always clear (otherwise I would not have attended). In each case I was invited because of my expertise as an infectious disease modeller and my role as co-chair of SPI-M-O. The exceptions were meetings organised by the SPI-M-O secretariat in which I had a role in setting the agenda and agreeing the minutes and consensus statements produced. I was asked by the SAGE secretariat

to comment on SAGE minutes occasionally when there was a technical aspect being recorded.

3.47. SPI-M-O is a group of quantitative scientists who specialise in the analysis of infectious disease transmission dynamics (see explainer in paragraphs 3.5 to 3.8). This area of science is key to understanding epidemics and making decisions about their management and mitigation. As a group we met over 100 times. Each meeting was recorded in minutes which were written by the secretariat and edited and approved by the co-chairs. Given the volume of work and the different areas that SPI-M-O covered, sub-groups of SPI-M-O were instituted early to focus on specific model types (see Explainer for policy-relevant classification of models).

3.48. SPI-M pre-pandemic modelling summary does not describe the scope of work that SPI-M-O would be responsible for. I had implicitly assumed that more statistical and established tasks, such as short-term forecasting, would be done by others, especially if they formed a formal government position within management of the epidemic. I expected that SPI-M-O would be mostly involved in supporting the technical functions of government and boosting the modelling capacity and creating the breadth of models required to generate ensembles to provide SAGE with scientific evidence.

3.49. In the event, SPI-M-O was solely responsible for the majority of the formal government modelling during the epidemic. The official SAGE website has 438 files containing 415Mb of data that were generated by SPI-M-O. The pre-prints and publications produced by members, and the large amount of work that was generated for SPI-M-O but did not get shared with SAGE will easily double the volume of work. The modelling done to support government done by the modelling groups associated with SPI-M-O was huge.

3.50. The question of what SPI-M-O should be doing, and how to prioritise the work and manage the resources available formed much of the work of the secretariat and policy co-chair. The appointment of Dame Angela McLean FRS (who was CSA in Department of Defence) as policy co-chair at the end of March was a very significant step and greatly improved the interactions between SPI-M-O members and government. The principal improvements were better understanding of what modelling could and could not do within government, and the consequent improved efficiency of what SPI-M-O did. In particular, modelling requires some policy

direction, as models must make assumptions about policy decisions (see paragraphs 8.43 to 8.67 and paragraphs 11.1 to 11.8).

3.51. The work that SPI-M-O completed falls into seven categories and I briefly describe each of these in turn. The explainer describes the key differences between some of these outputs from a policy perspective as I understand them, although note that my understanding has changed much during the epidemic.

Nowcasts

3.52. Assessment of the early stages of an epidemic relies largely on estimates of two quantities: the numbers of infections and the rate at which they are growing. As it is not possible to observe infections directly, the consequences of infection (test results, hospitalisations, deaths) must be used as indirect measures. The reproduction number, *R*, holds a central role in the theory of infectious disease dynamics (IDD), and SPI-M-O focused on that quantity, although this was perhaps inadvisable given the central role that the estimates came to have (for example, in the May 2020 strategy document *Our plan to rebuild: The UK Government's COVID-19 recovery strategy* [GM/30A – INQ000260628]). The directly observed quantity that can be estimated is the growth rate, and this might have been a better metric to focus on (see section 9).

3.53. SPI-M-O were producing estimates of *R* as soon as data were available, although not formally quantifying and combining the estimates – for example the consensus statement (27 April 2020) which went to SAGE 29 (28 April 2020) [GM/31 - INQ000223519] [GM/32 - INQ000053212]. In SPI-M-O consensus (4 May 2020) the underlying data for *R* estimates was presented but not formally combined [GM/33 - INQ000213253]. The first formal combination was in the SPI-M-O consensus (27 May 2020) which went to SAGE 39 (28 May 2020) [GM/34 - INQ000213254] [GM/35 - INQ000119951]. These combinations were generated through the same software as the short-term forecasts (see below). The basic methodology did not change although the process was modified up to July 2021 when UKHSA assumed responsibility for their production and release. Until then, SPI-M-O, through SAGE, produced official government *R* estimate every week.

Short-Term Forecasts

3.54. Management of the epidemic and its consequences, especially for health-care provision, is helped enormously by quantitative predictions of requirements. Accurate, quantitative predictions are only available in the short-term. Such models only forecast the data, so are only as meaningful as the data available. See Section 7 for more discussion about the availability and quality of data early in the epidemic.

3.55. I received an email dated 17 March 2020 with an urgent request for short-term forecasts from GCSA [GM/36 - INQ000213256]. Two groups responded to me the same day, and one soon enough to be sent to GCSA, and Imperial College also submitted directly to GCSA as requested [GM/37 -INQ000213257]. However, the data were of poor quality at that stage, and the forecasts therefore not very reliable. SPI-M-O had not had a chance to develop a consensus view, so these were well below the standard that I would have wished.

3.56. By 20 March 2020 groups within SPI-M-O had a series of models available, and developed a process to combine them. SPI-M-O needed guidance on what exactly was required in terms of the period of forecasts, frequency of forecasts and what was to be forecast which was not decided by 23 March 2020 [GM/38 - INQ000213258]. Interpretation of data remained an issue on 27 March 2020 and data from the devolved administrations was absent [GM/39 - INQ000213259] (see email thread).

3.57. By 27 March 2020 the technology for forecast combination was largely developed by Dstl with academic support from SPI-M-O members [GM/40 - INQ000213261]. I was very keen that if the nowcasts and forecasts being developed by SPI-M-O were to be official government statistics that government should be responsible for their final production. There was some discussion over the coming weeks about whether academic groups should be responsible for the combinations. I did not want academic research groups to be committed to producing these unless they were paid under contract (as I presume that Dstl were). Given that UKHSA did not take control of the process of combining forecasts until July 2021 I believe that this was the right decision. The level of quality assurance and robustness of production (to cover for periods of staff leave etc) required for official government statistics is outside of most academic groups capacity, but introduces a slower pace of work

that frustrated some members of SPI-M-O especially at the start of the process when they felt under pressure to produce.

3.58. Papers were sent summarising the proposed processes to SAGE 22 (2 April 2020), SAGE 23 (7 April 2020) and SAGE 24 (9 April 2020) [GM/41 - INQ000221949] [GM/42- INQ000075779] [GM/43 -INQ000120504]. The process had been largely set by SAGE 25 (14 April 2020) [GM/44 - INQ000120505] although they continued to be labelled "pilot". The devolved administrations were included by SAGE 27 (21 April 2020) [GM/45 - INQ000062295]. At some point over the summer 2020, it became clear that organisations using the SPI-M-O forecasts for healthcare outputs had developed their own methodologies and SPI-M-O ceased to produce them, but continued producing short-term predictions of numbers of deaths. Short-term forecasts continued until SAGE 51 (13 August 2020) [GM/46 - INQ000120551]. My understanding is that the data underlying the figures sent to SAGE were available throughout government to support immediate planning, but they contain very little information in terms of supporting strategic decision-making.

Medium-term Projections

3.59. I sent an email 13 August 2020 to SPI-M-O members that laid out my understanding of what was needed and suggesting that medium-term projections (MTP) would be useful. MTP fall between formal predictions and scenarios as they require the relatively strong assumption that transmission rates will continue to be the same as they have been in the recent past. School term dates were also included as a driver of behaviour with different groups making their own assumptions about what impact they would have [GM/47 - INQ000213268]. There was progress by 28 August 2020 – see email from SPI-M-O secretariat summarising discussions from a meeting the previous day [GM/48 - INQ000213269].

3.60. The first MTP was submitted to SAGE as a pilot for SAGE 56 (10 September 2020) and presented in the SPI-M-O consensus (18 September for SAGE 57 17 September 2020) [GM/49 - INQ000120554] [GM/50- **INQ000223534**] [GM/51- INQ000120558]. A limited number of groups were supplying data and there was no attempt to combine projections. The feedback suggested that these were useful for policy, and conclusions drawn from them were included in the SPI-M-O consensus (30 September 2020) at SAGE 60 (1 October 2020) with more groups contributing but no attempt at combination [GM/52 - INQ000213274] [GM/53 - INQ000120560].

3.61. An 'explainer' was published 31 October 2020 describing the process of producing the MTP and outlining the models used by the different groups [GM/54 - INQ000213276]. To the best of my knowledge, projections such as these had not been done previously, and they need to continue to be developed. However, they were very useful for understanding the impact of current transmission patterns, and I believe that they were useful for policy and managing the epidemic. The responsibility for producing MTP was handed to UKHSA in July 2021, and they are still being produced in May 2023.

Medium-term Projection Scenarios

3.62. The MTP were developed during October 2020 into medium-term projection scenarios (MTPS). The MTP are scenarios in the sense that they demonstrate the expected future 'if R remains unchanged'. The MTPS take this a step further and add 'if *R* changes to different values'. From a modelling perspective MTP and MTPS are very similar, but from a policy viewpoint they represent the impact of alternative measures. The *R* values used for different projections was agreed in discussion between members and secretariat, so contained a mixture of what modellers guessed might happen and what policy would like to compare. The models worked by generating patterns of human behaviour that generated the desired *R* value, and occasionally the *R* value that had been suggested was very difficult to get out of the models.

3.63. The first MTPS was included in the SPI-M-O consensus (21 October 2020) for SAGE 63 (22 October 2020) which showed the potential future trajectories if *R* was 0.6 compared to *R* remaining at its then value of 1.2 – 1.4. SPI-M-O continued to produce MTPS throughout the epidemic [GM/55 - INQ000213277] [GM/56 - INQ000087467]. MTPS require a date at which *R* value will change, and they were most useful when there was a policy decision to be made at a specific point, such as during the period January 2021 – July 2021 when the decision points were known well in advance.

Scenarios

3.64. Scenario modelling is the basis of modelling evidence for developing the strategy for the whole epidemic, or for longer periods, especially at the start of the epidemic when the data for other forecasts are not available. Most of the questions the Inquiry have asked are about scenario modelling. Generally, I believe that scenario

modelling was underused January 2020-December 2020, but used correctly after that. This is further discussed in paragraphs 4.26, 6.15, 8.43 to 8.53, 8.67 and 10.8.

Specific interventions

3.65. SPI-M-O were asked many questions about specific measures and interventions. I have not provided an exhaustive list. During the first months of the epidemic, we were asked about the impact of individual and household quarantine and school closures. As testing capacity increased in April and May 2020, SPI-M-O produced work on contact tracing. SPI-M-O was also asked to model bubbles, segmentation, local (spatially targeted) interventions, impact of festive periods and various testing strategies. All are addressed in the SPI-M-O consensus statements and supporting documents.

3.66. The 'ready-reckoners' that SPI-M-O members produced I think had an important role in helping policymakers understand the potential impact of changes in regulations. The SPI-M-O policy co-chair was instrumental in supporting their development and taking them into government. They were first presented at SAGE 22 (2 April 2020) [GM/41 - INQ000221949] [GM/57 - INQ000213279] and showed the impact of different policy options on the reproduction number as the adherence to social contact reduction. The authors were able to include vaccination and they were presented to SAGE 86 (8 April 2021) with several iterations in between [GM/58- INQ000213280].

Characterisation of Variants

3.67. The appearance of the alpha variant in November 2020 was a surprise to me as I had been implicitly expecting evolution to be gradual rather than jumping from one variant to another. The delta variant (March 2021) and omicron variants (November 2021) were similar in that the new variant replaced the circulating variants. At some point I expected that there would be a variant able to reinfect, so the omicron variant was less of a surprise. Some variants appeared in the UK but did not become established, e.g., the beta variant. In each case, SPI-M-O were asked what impact the new variant would have on the epidemic and the interventions that were in place.

3.68. This represented a huge amount of work, most of it was data analysis to estimate the important epidemiological parameters of the new variant then scenario modelling to understand the implications. All of this work is reported as documents

to SAGE and in the SPI-M-O consensuses. The ability to link data streams between individual patients proved invaluable for the data analysis. However, it was lucky that the different variants had different testing characteristics, in particular the presence of the S-gene target in the PCR tests used in the UK alternated between variants. Consequently, it was possible to track the rise of the alpha variant in terms of the proportion of tests that were S-gene negative, and the rise of the delta variant in the proportion that were S-gene positive rather than having to wait for sequencing data. Omicron was S-gene negative again. Without this alternation, understanding the relative transmission advantage of the different variants would have been much more difficult.

Explainer: Different types of models and model outputs

3.69. Models can be classified in many ways. Many of these are technical (e.g., stochastic vs deterministic, individual-based vs differential equations) but the key aspect for policy development is the balance between data and assumption. At one extreme, statistical models are data-driven, with few structural assumptions, and many of the assumptions made can be tested for validity. At the other extreme, (mathematical or theoretical) models can be primarily assumption driven. In lay terms, statistical models can be used to answer questions such as 'given current trends, this is what is likely to happen', and more assumption-driven models to address 'what-if' questions. Especially important for assumption-driven models is that they can be used to generate *counterfactual* scenarios, answering questions about how things might have happened differently, for example, 'what if UK had locked down 30 March 2020 instead?'.

3.70. The outputs of the two extremes are also different. Statistical models can be used to generate *predictions*. I use this word to mean proper quantitative forecasts, which provide an expected probability distribution within which the future observations are expected to fall. The models are testable (i.e., validated) by comparison between the prediction and the observations. If the model fails to predict the future, then either there has been a change in the data, or the assumptions used to generate the model were wrong, but note that if the model does accurately predict it does not prove the model correct.

3.71. In IDD modelling, predictions are only possible short-term. *Nowcasts* are formal predictions of what the current state of the epidemic is. Because of the inevitable

delays in the data, it is impossible to estimate directly what the current state is. The process of nowcasting is to analyse the data up to the point in the past it is known and then forecast it to the present. If the data (and particularly the reporting delays) can be well characterised, and there are relatively few assumptions, these estimates can be considered predictions in the strict sense. Nowcasting typically involves a single data stream (e.g., deaths), but nowcasting the epidemic should be some combination of different data streams, and this is essential if the reporting delays in different data streams are liable to change.

3.72. Beyond now, forecasts can provide predictions because of the delays in progression from infection to disease. For COVID-19, the delay between infection and death is about 3 weeks, so that the number of deaths over the next 3 weeks are largely determined by the number already infected. Consequently, a nowcast of the numbers infected implicitly provide a prediction of numbers of deaths. Providing forecasts of the numbers of deaths beyond 3 weeks requires being able to know the numbers that will be infected in the future, and since this depends on human behaviour, is not generally predictable. Outside of epidemics, average human behaviour is relatively constant (see explainer at paragraph 2.9 to 2.12) so that longer-term predictions are possible. However, given that human behaviour (and therefore contact patterns and therefore infection rates) within an epidemic are unknown, formal predictions beyond 3 weeks are not possible.

3.73. Assumption-driven models generate *scenarios*. Scenarios are based on a large number of assumptions and are not formal predictions. Comparing the future against scenarios does not reveal any 'truth' about the model, given that at least one of the assumptions is likely to be wrong, especially if they are quantitative assumptions (e.g., that the contact rate between people decreases by 50%). Models used to generate scenarios should be fitted to the past data in order to demonstrate that they are capturing the important features of the epidemic, but a model that fits the past well is no guarantee that it will predict the future. The principal assumptions are about future human behaviour, and in the case of the COVID-19 epidemic, about future government policy. For example, if modellers do not know when the next policy decision is to be made or what it will be, then they must guess at when and how behaviour might change, and even if modellers do know about future policy, they need to guess the specific impact.

3.74. In order to extract policy-relevant information from scenario models, it is necessary to consider multiple scenarios, ideally from multiple models (see paragraphs 8.1 to 8.10). This comparison can provide insight into future, qualitative patterns, for example that all the scenarios show an epidemic in the next two months. Sometimes it is possible to provide quantitative insight, for example if no model in any scenario generates a peak of hospitalisations greater than 1000 per day, then it should be bought to the attention of policymakers. However, this is not a formal prediction so cannot be formally assigned a likelihood of occurring. It is not usually possible to give different scenarios prior probabilities of occurrence (see paragraphs 11.12 to 11.22).

3.75. The principal outcome of comparison of scenario models is the degree of uncertainty in the future, not the expected, or average, outcome. Comparison between the scenarios will also generate insight into the critical drivers of the uncertainty. See paragraphs 11.15 to 11.20 for more discussion of the role of uncertainty in decision-making. ***[explainer ends]***

3.76. From my perspective, the purpose of the paper prepared by SPI-M-O co-chairs on 20 April 2020 was to facilitate policymakers thinking about the long-term scenarios, and how short-term patterns and goals fed into different policy choices [GM/58A – **INQ000148837**]. It was also a 'menu' of activities that SPI-M-O could be involved in from an epidemiological viewpoint. I believe that this document fed into discussion about what SAGE and SPI-M-O would be asked to do, but I did not see the document again.

3.77. In terms of tactical questions that SPI-M-O provided evidence for, as explained in paragraphs 3.47 to 3.68, given the differences model outputs, short-term forecasts are suitable for planning and management of resources required during the epidemic. Scenario models are suitable for informing strategy. MTP fall between the two and NHS were using them as forecasts and policy using them to support decisions. SPI-M-O were also asked about specific interventions.

3.78. In paragraphs 8.43 to 8.67 I discuss more of the role of SPI-M-O in developing evidence for strategic questions and policy choices.

3.79. The Inquiry asks for more clarification regarding my writing that: "There has been some misunderstanding and misrepresentation of the SPI-M-O role" in a recent paper [GM/59 - INQ000213281]. I was thinking of the public, and the steps I took

to resolve this was to answer questions to the media. I also supported the SPI-M-O and SAGE secretariat communications teams to write explainers that were made public. Within government, the SAGE and SPI-M-O secretariats and regular SAGE attendees fully understood SPI-M-O's role. I believe that the members of SPI-M-O knew what our role was.

3.80. The observations that SPI-M-O was not asked to consider beyond the transmission dynamics, nor had the capacity to do so, are correct. The terms of reference state in paragraph 8: "Questions of the operational practicality, proportionality and value for money of policy options are outside of the Group's remit." When members and I raised questions about the scope of the committee, the secretariat informed us that evidence on the 'cost-effectiveness' of interventions was not required. It is explicitly stated in SAGE 58 21 September 2020 minutes (para. 5) that the economic harms of interventions were being addressed outside of the SAGE structures [GM/60 - INQ000212102].

3.81. There are several issues that follow from this observation. First, it is not clear that this is a failing. At some point the output from modelling infection dynamics has to be combined with an understanding of the economic, educational and wider social impacts, but, in my view, trying to do this on a single committee that was simultaneously addressing modelling would be cumbersome and potentially degrade the quality of the modelling that could be done. The decision-makers have to balance the various harms done by the epidemic and the interventions, including the quality of the evidence provided. Pulling together modelling results with all the other considerations, and deciding the quality of that evidence would have been a large, multi-layered, multi-disciplinary enterprise that it not best done by SPI-M-O. I was assured that the quantitative exploration of the impact of measures on the economy, education, mental health and societal well-being was being done. In any case, there is good argument that data from different disciplinary expertise is better combined by people independent of the generation of the evidence. The modelling evidence was combined successfully with clinical and virological spheres on SAGE.

3.82. Second, is a question of capacity. The models the members were using were not designed for evaluating the cost-effectiveness of different NPI interventions, which is a relatively large undertaking and requires the input of other expertise, especially economists and health-economists. If this had been the required output from SPI-M-O, then the committee would have been composed of different members.

3.83. A third, and opposing, view is that although producing multidimensional output from SPI-M-O would not be optimal, having the context for the models is very important, given that there is no such thing as a policy-neutral model. Having some indication from policymakers what a better outcome would look like is essential to construction of useful modelling evidence. It is up to decision-makers to define what is meant by 'better outcome', but the modelling can certainly provide scenarios based on the different policy choices that are considered feasible.

3.84. Although I had some influence with the SPI-M-O secretariat and policy co-chair, the agenda for SPI-M-O was set by policymakers and decision-makers. I was not able to determine the scope of the agenda. The members raised concerns about the lack of policy context for modelling (see next paragraph), and especially which variables would define a better outcome. This was raised many times on SPI-M-O and SAGE, but ultimately it is for government to determine which evidence it considers for which purpose. A key part of the decision-making role is deciding what is important and required for making the decision.

3.85. During the summer and early autumn 2020, there was interest amongst some members of SPI-M-O to develop an 'objective function' that could be used to judge between different outcomes (see Explainer below). It was discussed at SPI-M-O main meeting 23rd September 2020 and other meetings [GM/61 - INQ000213284]. I received an email which suggested that observers had found it useful and were taking it up. I forwarded the email to the SPI-M-O secretariat and advised the SPI-M-O to whom it had been sent to continue to engage, but keep the secretariat informed. I heard nothing more about this [GM/62 - INQ000213285] [GM/63 - INQ000213287].

3.86. I do not have any direct knowledge of the analysis being completed within government on the wider economic considerations of decisions. There was some discussion during the epidemic (for example, [GM/64 - INQ000213288]) but I and SPI-M-O were not directly involved or consulted.

3.87. Without policymakers/decision-makers being able or willing to share even a generalised view of what a good outcome would be, we could not make progress. This is not necessarily a criticism, but if in future epidemics SPI-M-O is to be tasked with addressing the wider impact, then some guidance will be required as to how the different impacts are to be traded-off against each other, i.e. what is a 'better' outcome,

although given that epidemics are inevitably damaging this is better framed as a 'least worse' outcome.

3.88. Overall, it is my belief that policymaking and decision-making in future epidemics would be considerably improved if there was more communication and feedback between groups addressing the different aspects of the impact of infection, interventions and their interactions.

3.89. I do not agree completely with the statement, "…led to advice being less robust than it could have been if a broader group of experts was empowered to provide analysis and advice from the outset" but I understand the issue.

3.90. Inevitably there are differences between academic groups working in similar areas. Funding competition means that groups must find their niche, and there are limited resources, so that groups will always tend to vary in size. Groups also differ in terms of their experience with analysis of epidemics in real time. Groups also differ in terms of their collaborations and contacts globally and consequently, if the epidemic emerges first outside of the UK, will have different access to initial data and 'on the ground' experience. This was the situation in the UK January 2020, and is almost certain to be the case in a future epidemic unless there is a revolution in funding mechanisms. Inevitably smaller groups have less funding and less flexibility. If government values a modelling committee producing consensus and ensemble modelling, then funding will need to be provided. There is more on this when discussing in paragraphs 3.32 to 3.41.

3.91. The key question is to what extent and how does this heterogeneity influence the evidence, and how can the negative consequences be mitigated. One of the criteria that I imposed on myself when recommending which people and groups should join SPI-M-O was that there would be a sufficient critical mass to ensure evidence would be as robust as possible. Given that I was not able to predict how many people would be able to contribute at any one time, nor how long SPI-M-O would be asked to sit, the number of contributors invited was relatively large. They joined as early as possible in February and March 2020. In the event, virtually all of those asked responded positively and contributed throughout the two years. I agree with the witness that this large number contributed greatly to the reliability and robustness of evidence.

3.92. The issue seems to be that at the start of the epidemic, the heterogeneity in group capability resulted in less robust evidence because only a smaller number of groups

were able to contribute. The groups from LSHTM and Imperial College were better connected (in terms of international and national standing), had more resource (including people) and were more experienced in analysis of epidemics in real time so they were able to contribute more, and more quickly. The delays in data provision hampered the involvement of other groups for whom this was their only data source. I understood the frustration that others felt at not being able to contribute as much as they would have liked, but there was nothing that I could do to rectify. Essentially, for the first few weeks of the epidemic, others were being asked to evaluate the work produced by the larger groups. Whether the evidence provided by SPI-M-O would have been more robust if other groups had been able to contribute to the same level is not obvious, although I would have been more confident if there were three groups (rather than two) operating at the same level. However, with hindsight, I do not believe that any of the advice coming from SPI-M-O was "wrong", and that the problems with providing evidence derived more from availability of data and other resources, the confusion about strategy, and what we were being asked to do.

3.93. In order to correct this imbalance at the start of the next epidemic, there needs to be a revolution in funding mechanisms and strategic allocation of resources. The foundation of UKHSA and the Centre for Pandemic Preparedness (announced June 2021 [GM/65- INQ000213289]) has already changed the landscape considerably.

3.94. The witness is correct that I was asked "on multiple occasions by various members," about inclusion of economic criteria into models, and that my response was that "economic evaluations go beyond the remit of SPI-M, and that social and economic aspects would be accounted for in SAGE, or directly in Government." This is discussed earlier in this section in paragraphs 3.80 to 3.86.

Explainer: Models outputs and use

3.95. Models can produce any output. The basis of a model of transmission dynamics has to be infections, and typically modellers will focus on the outcome of infections as the principal output of the model. In the case of COVID-19 this was hospitalisations and deaths, although infections were equally important prior to vaccination rollout. If the model includes specific interventions, then these can be output as well, and a typical use of a model would be to evaluate a vaccination programme as numbers of deaths averted per vaccine dose given. Using this approach models can be used to determine the most efficient use of vaccines, or any other intervention. Consequently, it is at least

theoretically possible to determine during the COVID-19 epidemic the number of deaths averted per lockdown day and see what impact different intervention patterns would have. Work started during summer 2020 has subsequently been published, but was not used during the epidemic [GM/66 - INQ000213290].

3.96. Models can be used to explore different strategies, and to find an optimum if the outcome is defined. The outcome is known as the 'objective function', for example, the number of deaths averted per vaccine dose given. Objective functions can be complicated and have multiple criteria. Deciding the objective is a necessary step before exploring different strategies, although they are constrained by what is possible to achieve given the biological properties of the problem. ***[explainer ends]***

3.97. I did not organise any meetings which government officials used for gathering and understanding scientific evidence. I was invited to such meetings by government officials and I did not keep any record of the membership or attendees. The Inquiry asks specifically about the sub-groups of SAGE which I describe below

3.98. Enduring prevalence sub-group. This group started meeting in March 2021 although there had been discussion about the geographical heterogeneity on the spatial epidemiology sub-group of SPI-M-O for some months. I believe that the idea of this sub-group (originally called the 'correlates of local heterogeneity' group) was to examine the drivers of local prevalence. There was an indication that working from home (WFH) was an important driver and this was raised at the start as in this email 25 March 2021 [GM/67 - INQ000213291]. I do not believe that anything particularly useful came of this suggestion, partly because the areas used by ONS were not consistent with the areas used for reporting cases. The modelling input was largely based on data analysis of spatial patterns as described in email dated 18 April 2020 [GM/68 - INQ000213292]. The group produced a report which went to SAGE 87 (22 April 2021) [GM/69 - INQ000213294] [GM/70 - INQ000119963]. With hindsight, this group lacked a specific question to address, and much of the discussion was 'problem framing'. At the same time, JBC were getting more analytical capability and were able to do more of the analysis that SPI-M-O was doing. Policy questions were focused more nationally on ensuring that a vaccination programme was rolled out. The group continued to meet until March 2022.

3.99. Care homes sub-group. Care homes (or long-term care facilities, LTCF) were considered generically during February and March 2020. The first record I have of a

formal consideration was 17 April 2020 with an email discussion with PHE [GM/71 - INQ000213297], and subsequent discussion on SPI-M-O (20 April 2020) as recorded in consensus [GM/72 - INQ000213298]. The SAGE sub-group was set up 24 April 2020. My role was largely to be sure that SPI-M-O member Hall was comfortable chairing and co-chairing, and I played a very minor role.

3.100. There seemed to be very little data coming from the LTCF sector initially. It is a complicated structure with mixtures of private/public provision and reporting frameworks clearly not set up for rapid, national emergencies. SPI-M-O member Hall played a large role in co-ordinating and developing reporting and analysis framework. The work of the sub-group was particular focused on the impact of testing strategies on the risk of outbreaks within individual LTCF, and on reducing the impact when it occurs. They also highlighted the potential roles of domiciliary carers and informal carers – see paragraphs 8.27 to 8.40 for more discussion of the role of LTCF in the overall epidemic.

3.101. Schools/children sub-group. SPI-M-O was asked to consider school closures as a means to reduce transmission early in February 2020. This came to SAGE in terms of multiple modelling papers and SPI-M-O consensus documents before schools were closed 23 March 2020. SAGE instituted a special sub-group which reported to SAGE 30 (30 April 2020) [GM/73 - INQ000074907] [GM/74 - INQ000075781]. I played a relatively minor role in the sub-group.

3.102. Higher and further education sub-group. I was not very close to the analysis done by SPI-M-O members so cannot give any firm views or timelines for this work beyond the documents submitted to SAGE. I contributed to various documents, such as attached to this email [GM/75 - INQ000213301] [GM/76 - INQ000213303], although I do not believe that it went to SAGE but straight to the Department for Education.

3.103. Nosocomial transmission sub-group. This was set up by SAGE end of March 2020, although there had been discussion about the potential role of nosocomial transmission (i.e., transmission that happens in a healthcare setting) especially in hospitals. A previous novel virus (SARS, now SARS-CoV-1) was highly infectious within healthcare settings, but not very transmissible outside of hospitals, so there was some intuition on SPI-M-O that hospitals could be a focus.

3.104. The first indication that this might be a problem came from the FF100 dataset which showed a large number of infections with apparently no clear risk of transmission –

see email 15 March 2020 [GM/77 - INQ000213304]. However, there was considerable uncertainty in the data at that point (see paragraphs 7.6 to 7.21). My concern was that if the explosive growth was real, and had been discovered since surveillance in ICU had been implemented then it could be uncovering significant nosocomial transmission – email dated 16 March 2020 suggesting to PHE that data were required to address this hypothesis [GM/78 - INQ000213305]. This was discussed at SPI-M-O and included in the SPI-M-O consensus statement (20 March 2020, paragraph 7 [GM/79 - INQ000213306].)

3.105. During April 2020 I was concerned that, given normal contact rates had been reduced so that overall the epidemic was decreasing, that the remaining significant transmission was associated with health-care, and particularly hospital, settings as in this email conversation with GCSA dated 20 April 2020 [GM/80 - INQ000213309], and this email relating to the situation with LTCF 16 May 2020 [GM/81 - INQ000213310]. The SPI-M-O consensus at SAGE 27 (21 April 2020) [GM/45 - INQ000062295] was that nosocomial transmission was likely a significant driver of on-going transmission: "CO-CIN and PHE modelling is consistent with 10-25% of hospital confirmed cases in England being acquired in hospital. However, this varies between hospitals and is likely an under-estimate of the total healthcare acquired infections." In particular, the modelling was not able to include transmissions from infections acquired within hospitals. Data developing from LTCF suggested there might be three linked epidemics to consider as in email 1 May 2020 [GM/82 - INQ000213311].

3.106. Modelling analyses developed during April 2020. LSHTM analysis estimating the rates of infection based on community and nosocomial transmission SAGE 27 (21 April 2020) [GM/83 - INQ000213312]. SPI-M-O member Read, working with CO-CIN PI Semple, produced an analysis of the CO-CIN data which showed that there were considerable differences between NHS Trusts in England. The response of NHSE was that the analysis was insufficient for assessment or management, which I disagreed with as in the email 29 April 2020 [GM/84 - INQ000213313]. This analysis has since been published and shows that whilst there was a significant amount of transmission in hospitals, it was not as dominant as it could have been given the data at the time [GM/85 - INQ000213314].

3.107. Mass testing sub-group. Mass testing was considered by SPI-M-O members in April 2020 with some preliminary analyses going to SAGE 24 (9 April 2020) [GM/43 - INQ000120504] [GM/86 - INQ000213315]. There was no follow-up as far as I am

aware until August 2020 when SAGE asked for a focused, interdisciplinary consideration reported to SAGE 53 (27 August 2020) [GM/87 - INQ000213316] [GM/12 - INQ000213048] . There was further consideration of mass testing as an intervention in November 2020. The consideration appeared to me to be rushed given that it had been proposed / announced in early September 2020. See for example the email thread of 9 November 2020 [GM/88 - INQ000213317].

3.108. Testing itself does not achieve any control but provides information for targeted interventions – a LD of the whole population is equivalent to everybody having tested positive, compared to a 'lockdown' of people testing positive. There seemed to be relatively little consideration of the consequences of testing rather than the implementation of testing. In any intervention that is targeted on the basis of testing, the compliance with the interventions for people who test positive is always the major factor in determining the impact and cost-effectiveness. If mass testing/screening is to be a potential intervention in future epidemics, then there needs to be concerted consideration prior to the next epidemic. Mass testing/screening has the potential to be an important part of an overall strategy, but is unlikely to be effective as a single measure, and needs to be carefully implemented.

3.109. Vaccine sub-group. This was a scientific discussion group. The group reported back to SAGE, largely about research approaches to improving and understanding the characteristics of vaccines rather than providing evidence for infection control policy. The decision was taken as the vaccination programme was rolled out to give the vaccination programme primacy in deciding policy, i.e., that any NPI evidence was 'given' the vaccination programme. The output from the sub-group can be read in the documents produced.

Section 4: The Early Stages of the Pandemic

4.1. The activities of SPI-M-O during January – March 2020 have also been covered in other sections of the statement, especially sections 3, 5 and 7. All the work that SPI-M-O did is reported in the consensus statements and minutes of the meetings. I am happy to answer specific questions beyond those below.

4.2. a. Preliminary estimates of the incubation period and mortality rates were made from overseas data during February 2020. The growth rate in the UK could only be

estimated once the epidemic had started and the data streams were reliable which was not until mid-March 2020 at least.

4.3. b. The weekly combined estimates of the growth rate and the reproduction number are as reported in the SPI-M-O and SAGE consensus statements.

4.4. c. Short and medium-term projections are addressed in paragraphs 3.47 to 3.68.

4.5. d. The SPI-M-O responses to policy-specific questions are reported in the consensus statements. Note that SPI-M-O was not asked about 'bubbles' until May 2020, and contact tracing until April 2020.

4.6. I was aware of the papers published at the end of January 2020 in *The Lancet* and *NEJM*. I, and I believe all other SAGE attendees, had read these papers. The findings of these papers, combined with other evidence, raised a high alarm within the SPI-M and SPI-M-O members, and these papers were likely one of the reasons that SAGE and SPI-M-O were instituted. No individual scientific paper is conclusive evidence, and the point of SAGE is discussion and consensus of the total evidence. Generally, SAGE did not pick through individual pieces of work. Occasionally, researchers were invited to present their work and answer questions on SAGE and SPI-M-O where there was a clear benefit to understanding and a clear relationship to policy. To some extent the point of sub-groups on SAGE is to go through original research so that SAGE can focus on the overall view. I cannot comment on the communication of the degree of alarm raised as a result of these papers to decision-makers. To my understanding and knowledge, there was no concerted attempt to communicate risk to the public at that time.

4.7. I did not have direct discussion with David O'Connor. The email chain is about the need for data from other countries to inform the outcome in the UK and provide parameters for models [GM/89 - INQ000213318] [GM/90 -INQ000213320]. All my interactions with government officials were cc-ed to the SPI-M-O secretariat, although more frequently I interacted with the secretariat who then communicated directly within government.

4.8. My role generally was to explain to those in government what data were required to address the questions that government were asking. There was great interest in gaining accurate detailed data from countries that were further along in their epidemics than UK, and China was first in that list. The email chain demonstrates this interest,

not just from SPI-M-O, but also to inform public health interventions (such as isolation time).

4.9. At SAGE 11 27 February 2020 the key parameters that would determine the level of health-care demand (and particularly hospital care) were discussed. Throughout February 2020 it became increasingly clear that NHS capacity in the UK would be overwhelmed. SAGE asked that a working group be set up jointly with NHS to discuss the extent of the overwhelm and what the detailed pattern of resources could likely be, and that the results of this meeting be reviewed by SPI-M-O (SAGE 11 minutes second bullet after paragraph 9 [GM/91 - INQ000075777]).

4.10. In order to expedite the process the working group met on Sunday 1 March 2020. I did not attend, but SPI-M-O secretariat did, and kept me informed [GM/92 - INQ000213322]. The outcome from the meeting was discussed at SPI-M-O (2 March 2020) and is presented in the SPI-M-O consensus statement which was discussed at SAGE 12 (3 March 2020) [GM/93 -INQ000213324] [GM/94 - INQ000213325] [GM/95 - INQ000119719].

4.11. The paper entitled *Adoption and impact of non-pharmaceutical interventions for COVID-19* presented to SAGE 12 (3 March 2020) is essentially a review and summary of the understanding of the epidemic potential in the UK and the potential impact of different NPI at the time [GM/95 - INQ000119719][GM/96 - INQ000213327]. I was not involved in production of this paper, but do not disagree with it in any way. It was not a SPI-M-O paper, but prepared by two principal modelling teams (Imperial College and LSHTM) and presented by the two independent SAGE attendees who led the teams.

4.12. The paper lays out two main areas. First, a summary of the current understanding (summarised by the reproduction number) of the epidemic in other countries. Second, the interventions that had been implemented in other countries with a summary of their potential impact. There was very little surveillance data available at the time so that there was much uncertainty, and the paper largely lays out what is still unknown, but which is expected to be known in the coming weeks.

4.13. This, and other work, led to the production of the subsequent papers [GM/97 - INQ000194008] [GM/98 - INQ000213329]. On the 3 March there was perhaps still a faint hope that the UK would somehow avoid a large epidemic, although this had dissipated by 9 March. In any case, a sensible basis on which to prepare strategy would be that there would be a large epidemic in the UK. The papers presented to SPI-

M-O on 2 March 2020 – some of which were included in SAGE papers, but all summarised in the SPI-M-O consensus paper to SAGE 12 (3 March 2020) – give a fuller idea of the understanding at the time [GM/94 - INQ000213325] [GM/95 - INQ000119719].

4.14. I cannot recall the details of the discussion on 3 March surrounding this paper, partly because it was one of many, and quickly developed into the subsequent papers.

4.15. Herd immunity is not a strategy. It is the inevitable outcome of an epidemic. In the interview I gave 13 March 2020, I was not advocating a strategy. I was attempting to convey the idea to the public that the epidemic would be a major event, but that it would end at some point, that this was going to be a very memorable period, but only a period.

4.16. For a fuller understanding of my views in giving this interview, please also see the comment published 22 October 2020 in the Lancet [GM/99 - INQ000213330].

4.17. I am still unsure what a "herd immunity strategy" is. In the same interview I said that an uncontrolled epidemic would be the worst outcome in my opinion. The principal public health challenge would be to manage the epidemic to do as little harm as possible. The principal political challenge would be to avoid a catastrophe. Whichever strategy, tactics and implementation would be used, the outcome would be herd immunity. In the event, the development and deployment of vaccines greatly reduced the morbidity and mortality resulting from the epidemic, so that the UK is approaching herd immunity now (May 2023) with considerably less morbidity and mortality than we would otherwise have had to endure.

4.18. There is a good analogy between an epidemic and falling because of gravity. When somebody is falling, the end point is reaching the ground. How you get there and what is there (parachute, trampoline etc) is a separate question. Gravity pulls you down until you stop. Epidemics continue until you reach herd immunity. It was perhaps naïve of me to try to explain the end of the epidemic before it had really started, but I was not advocating a strategy.

4.19. I am unsure why the government response at the time was termed 'herd immunity strategy', although I presume that it's related to myself and GCSA using the term 'herd immunity' in interviews coincidentally on the same day. I was not aware that GCSA had given an interview on the same day that I was interviewed.

Explainer: Herd immunity

4.20. Herd immunity is the term used to describe the epidemiological situation in infectious disease when there is sufficient immunity in the population, i.e., there are more than a critical proportion immune, such that the number susceptible is reduced below the threshold required for an epidemic. This is a natural phenomenon.

4.21. The incidence of infection in a population is essentially the product of three numbers:

4.21.1. the rate of effective contact between infectious and susceptible individuals,

4.21.2. the proportion that are susceptible and

4.21.3. the number of infected individuals.

'Effective contact' is determined by the route of transmission and how good the virus is at transmitting as well as the way that people are behaving. The number of infectious people increases during an epidemic, and there are only two ways to slow it, either changing behaviour (e.g., generically reducing contact rates, or isolating infectious people) or reducing the number susceptible (e.g., vaccination). Herd immunity occurs when the number of susceptible people is reduced sufficiently that a sustained exponential growth in infections is no longer possible.

4.22. Although herd immunity has a firm theoretical and phenomenological basis, there is still some imprecision in its formal definition. There is some conceptual overlap with the term endemic, in which, averaged over a long period, the reproduction number, *R*, averages to unity, in which case although there might be periodic epidemics, there is prevailing immunity. However, it is mostly used in relation to the implementation of vaccination programmes, where it is used to define a coverage required to prevent on-going transmission. The best example of the two uses is in the epidemiology of measles in the UK. Pre-vaccination there were periodic epidemics in which herd immunity is reached at the peak of each epidemic. Post-vaccination elimination of measles was achieved by vaccination at a sufficiently high level to prevent transmission.

4.23. Another nuance for herd immunity is that its impact depends on the amount of contact. If contact is low, then herd immunity is achieved more easily. If herd immunity is sufficient to curtail exponentially increasing transmission, but then the amount of

contact increases, herd immunity will no longer be sufficient and a new wave will occur. ***[explainer ends]***

4.24. The Inquiry asks about two separable issues here: the resources available for SPI-M-O to produce its modelling evidence, and the issue of lack of consideration of the whole epidemic at the start.

4.25. In terms of resources, I have addressed this in paragraphs 3.32 to 3.41 and paragraph 7.1 to 7.2 (iv). The "scrabbling" was for data, personnel and computational power. The impact was significant as detailed in those sections.

4.26. The failure to see the epidemic as a single entity and to decide the overall strategy has been addressed elsewhere. This can be remedied in future by deciding the overall strategy before the next epidemic starts. My belief is that the impact of the lack of overall strategy was huge and particularly felt during autumn 2020.

Section 5: The Timing of the First National Lockdown

5.1. The question of timing of interventions and their correctness is not a scientific question. A national lockdown has diverse impacts, of which changes in viral transmission is one. It is not the job of scientists to balance the different outcomes. The task given to SPI-M-O is to evaluate the impact of different interventions on transmission dynamics.

5.2. It is beyond question that all other things being equal earlier lockdown (hereinafter simply referred to as 'LD') would have reduced the number of infections (and hospitalisations and deaths) during the first wave. However, it's not clear that behaviour would have changed in the same way with an earlier LD, both during the first LD and subsequently. The change in transmission dynamics would have been reflected in the rest of the epidemic, particularly leading up to a second wave. The outcome of the first LD, if enacted earlier, is also likely to have changed the perception of the policymakers and decision-makers for subsequent decisions. These effects could be positive or negative.

5.3. Given that there was significant ongoing transmission within the healthcare and social care systems after the 23 March 2020, which was not halted by the LD, it is not straightforward to estimate the impact of an earlier introduction, even now, and doing so at the time carried huge uncertainty (see Explainer). This was principally because

of the state of the data flows (which meant that the then current situation was not clear) and the unknown policy, i.e., we did not know which NPI were to be included in a LD, nor did we know the date of such a LD.

5.4. Given all these caveats I am not sure that it's possible to say with any certainty what the optimum time point should have been without a clear set of strategic objectives being set beforehand. The last point is crucial – without some idea of what decision-makers regarded as a good, or least-worst, outcome, it is impossible to judge whether the first wave was too large or too small.

5.5. It is tempting to think that had LD been introduced 2 doublings earlier (so that the number infected was one quarter of what it would have been) that the numbers of hospitalisations and deaths would have been 25% of the observed totals. We now know that the transmission pattern within the UK was being greatly influenced by inflows from across Europe immediately prior to LD [GM/100 - INQ000213028]. Earlier introduction of travel restrictions would have prevented a proportion of this influx, and its subsequent exponential growth. At the same time there was an ongoing epidemic in healthcare and social care systems on which an earlier introduction would have had less impact than the general community [GM/101 - INQ000213029] . I am not aware of any detailed attempt to evaluate the impact of an earlier LD on the transmission dynamics.

5.6. SPI-M-O were asked about the timing of introduction of any measures (SAGE 12, 3 March 2020 [GM/95 - INQ000119719]). The LSHTM and Imperial College teams were the only groups to have useable, reliable models at that point. Without any indication of the scenarios to consider, the modellers essentially had to imagine what such a policy might contain. These groups produced scenarios that were later published [GM/102 - INQ000213030] full paper [GM/103 - INQ000213031] Imperial College report [GM/104 - INQ000213032]. Whilst very informative to those who understand models and modelling, the results being produced were considered too complicated to inform policy-makers, and were consolidated into a document that was discussed at SPI-M-O (9 March 2020) although it had been prepared 3 March 2020 (*GOS: Illustrative impact of behavioural and social interventions lasting several months on a RWCS*) [GM/105 - INQ000213033] [GM/98 - INQ000213329] and presented at the next SAGE [GM/106 - INQ000106152]. This summary paper was more easily understood. Note that the rapidity of the request and response meant that it was not possible to develop a formal SPI-M-O consensus on this document which partly

created the frustration amongst some members (see paragraphs 3.80 to 3.94). Dame Angela McLean FRS was closely involved in the creation of the summary document, and later became co-chair of SPI-M-O.

5.7. The question of whether SAGE should have advocated more strongly for particular interventions, or timing of interventions, is difficult to judge without hindsight. Personally, I think that SAGE and sub-groups should avoid overt advocacy, but should provide evidence to ensure that the decision is delineated for decision-makers as clearly as possible. Clearly, senior officials who might be on SAGE should be asked for advice based on the evidence or their experience, and they themselves might seek further opinions. The process of generating advice is very different from generating evidence and ideally kept as separate as possible. A key aspect of decision-making is selecting whose advice to seek and how to use that advice. Decision-makers are politicians, so the decisions they take are political, and therefore so is the advice. Evidence, especially relating to a natural phenomenon such as an epidemic, should be as impartial as possible and as transparent as possible. My understanding is that the main aim of SAGE is to provide evidence. There is not a clear boundary between evidence and advice, but if evidence-gathers advocate for particular policies, based only on a portion of the considerations that decision-makers have, then the distinction becomes more blurred. I am, of course, speaking from the perspective of a disciplinary sub-group of SAGE, which is more firmly established at the evidence end of the spectrum.

5.8. I tried throughout the epidemic to retain the difference between evidence and advice. One area in which I was challenged is the relationship between advice and evidence when the evidence is compelling for a particular policy choice. The evidence in February and early March 2020 accrued that an uncontrolled wave would have been devastating. The simple calculation available then was that >50% of the population (>33million) would have been infected with an infection fatality ratio of 1% would have resulted in >330,000 deaths and many more hospitalisations over a period of a few months (see paragraphs 4.9 to 4.10). In SAGE meetings 12 and 13 on 3 March and 5 March 2020 the figure of 80% attack rate (see Explainer below) was used to make the projected deaths 528,000 (0.01 * 0.8 * 66,000,000) [GM/95 - INQ000119719] [GM/106 - INQ000106152]. This evidence potentially acts as an imperative for action, and in this case the introduction of generalised NPI. My dilemma arose in terms of thinking that evidence should be public, but advice can be confidential. In the Newsnight

interview on 13 March 2020 I chose not to cite or try to explain the projected size of an unmitigated wave, but instead said it would be "…the worst outcome".

5.9. At the time that SPI-M-O were being asked about introduction of generalised NPI, we did not know what form they would have, nor what their impact would be in terms of reducing transmission. The only policy choice that we had good information about was the size of an unmitigated wave (see SAGE 12 3rd March 2020, *COVID-19 compared with NSRA pandemic influenza planning assumptions*) [GM/95 - INQ000119719] [GM/107 - INQ000213035]. An unmitigated wave would arise as a result of the policy choice to do nothing or very limited NPI, although this is further discussed in paragraphs 5.16 to 5.21. The uncertainty surrounding the measures that might be introduced at some time and in different places, and their effects, was a significant factor in shaping discussion on SAGE. My understanding at the time is summarised in a document amongst the SAGE 13 papers (5 March 2020, *Considerations for NPI policy – timing and sub-national targeting*) [GM/106 - INQ000106152] [GM/108 - INQ000213036]. This is discussed in detail in paragraphs 5.16 to 5.21.

5.10. The announcements 23 March 2020 were novel to me, and the peak of deaths (apparent by early April 2020) was a very welcome outcome but could not have been guaranteed on 23 March 2020. Generalised NPI introduced in other countries, especially China, reduced transmission to a large extent (see SAGE 10 minutes 25 February 2020 [GM/109 - INQ000087503]). The Italian LD was imposed on 8 March so was too early to be able to measure its effect. Personally, I felt that there were sufficient significant cultural, legal and political differences between China and UK that I could not assume that the UK would have a similar experience as China; however, this is not my area of expertise, and I did not voice this intuition at SAGE. Consequently, suggesting that SAGE advocate for particular action did not make sense, and still doesn't, if we do not have any certainty in what we are advocating for in terms of outcome. A key factor decision-makers have to consider in their decision is the uncertainty and it is critical that the analysts and advisors pass uncertainty to the decision-makers rather than remove it by highlighting the 'preferred policy choice' (see paragraphs 11.12 to 11.22).

5.11. Nonetheless, I felt considerable frustration that there was not a clear and apparent strategy to mitigate the epidemic wave that was impending and deal with the remainder of the epidemic. There was considerable anxiety amongst the attendees/members of SAGE and SPI-M-O to whom I spoke which gradually increased into early March. My

main concern was that government was not doing enough, either to prepare the population for an unmitigated wave or to introduce interventions earlier – the epidemic was developing faster than the response. I was also keen to have interventions introduced separately with the hope of being able to understand what the impact of each intervention would be. SPI-M-O had produced work on school closures [GM/110 - INQ000075778] [GM/111 - INQ000213040] which had less uncertainty (but still considerable) given that school closures had been used in previous epidemics. Had school closure been introduced separately to other measures there would have been considerable data to inform on their effects that would have been important information for decision-making in the remainder of the epidemic. This opportunity for learning was missed.

5.12. As well as being frustrated at the apparent inertia in introducing interventions, I greatly appreciated the care that DHSC was putting into develop isolation guidelines that were equitable and appropriate for all households. Generalised NPI had not been considered as a policy previously (see SPI-M Modelling Summary [GM/14 - INQ000212124]), so I expect there will have been very little preparatory work prior to 2020.

5.13. Consequently, in response to the direct question, I do not agree with the view that SAGE should have been blunter in calling for stronger action. The discussions on SAGE, I believe, imparted the concern, anxiety and frustrations, to the officials attending. But the quote seems to suggest that the group should have been advocating publicly which I do not think is appropriate.

5.14. The discussion about the impact of interventions on behaviour in both short and long term is not my area of expertise. The idea of 'behavioural fatigue' was expressed on SAGE, although I cannot recall where it first appeared. Note that I was aware that if the epidemic peaks were successfully kept low by behaviour change (mandated, suggested, or spontaneous) then the epidemic would have been extended and lasted, potentially, for years. There is evidence that general enthusiasm and adherence to NPI was waning by November 2020 in which a significant number of MPs voted against extending the government's ability to impose measures. Although it is not my field, I thought at the time that NPI would not be feasible indefinitely was a realistic concern, and that view has not been changed by the evidence. Generally, the main point was that government should not start doing something without thinking about its longer-term impacts.

Explainer: Attack rate and final size

5.15. The 'attack rate' is the technical term for the proportion infected. There has been considerable theoretical work on the total proportion of people who are infected during an epidemic, also known as the 'final size'. The final size of an epidemic is relatively complicated to work out, but the logic is the same as that for the 'herd immunity threshold'. Essentially, epidemics do not infect everybody because the transmission rate is reduced by contacts being increasingly people who are already infected or immune from prior infection. The final size is related to the basic reproduction number, R_0, so that the early estimates in the UK enabled some idea of the total number of infections in an uncontrolled epidemic. The closest we can come to seeing this in reality is in small communities that had exceptional behaviour. In one such study, the final size was 65% of the population [GM/112 - INQ000213041]. ***[explainer ends]***

5.16. I cannot comment on ministers' preferences. My underpinning concern up until 16 March 2020, and then 23 March 2020 was that the strategy for dealing with the whole epidemic was unclear. At the first SPI-M meeting to discuss the epidemic on 27 January 2020, the only guidance we had was that government wished to avoid having the epidemic during the winter, because of the likely healthcare overwhelm caused by simultaneous seasonal influenza and COVID-19 epidemics. Whilst this was very early in the epidemic, and knowledge accrued rapidly to make this a rather incongruous strategy with hindsight, I was not told differently throughout February 2020, and it reappears as an objective in a SAGE paper 4 March [GM/98 - INQ000213329]. To my interpretation, the announcements and public appearances of decision-makers indicated that they wished to have the epidemic as a single wave and did not envisage trying to prevent a wave with a Wuhan-like 'lockdown'.

5.17. Necessarily, any scientific evidence provided and advice given are framed by the strategic context. Without an explicit indication from decision-makers what they wish to achieve, or what would be the least-worst outcome from the epidemic the preferences must be inferred. Throughout February 2020, SPI-M-O was asked about the parameters that determine the scale of the problem – principally the reproduction number (which determines the attack rate), the infection fatality ratio (IFR) and the burden on clinical care (particularly the hospitalisation rates). The age-related risks of hospitalisation and death were apparent very early and a focus of attention. We were also asked about approaches to delaying the epidemic (see SAGE 4 4 February 2020, *SPI-M-O: Consensus view on the impact of possible interventions to delay the spread*

of a UK outbreak of 2019-nCov) [GM/113 - INQ000051925] [GM/114 - INQ000213043]. Seasonality was an early question, with the possibility that SARS-CoV-2 would not be transmitted in warmer weather, although SPI-M-O was always sceptical of this [GM/115 - INQ000075784] [GM/116 - INQ000213045]. We were also asked about the impact of school closures and although our views refined as more information and results were available, there was not a huge change from the first consensus view [GM/115 - INQ000075784] [GM/117 - INQ000213046]. The questions in terms of managing the epidemic were related to delay which later became known publicly as 'flattening the curve' (e.g. see SAGE 7 13 February 2020, List of Actions in minutes [GM/118 - INQ000106109]). The impact of more generalised interventions on the whole epidemic was not being considered until later in February [GM/109 - INQ000087503] [GM/91 - INQ000075777]. None of these tasks addressed the longer-term strategy if the impending epidemic had been stopped. In particular, if measures are introduced to stop the epidemic, then the epidemic will restart when the measures are sufficiently relaxed so that the country is in the same position again at a later date.

5.18. Against this backdrop, it was clear to me that the government's preferred strategy was to have the epidemic in a single wave. At the time, I did not have a strong view in the sense that the measures that would have to be introduced to prevent the epidemic were untested but would have a significant negative impact themselves. I was not taking the decision on how to balance the different harms. However, it was becoming clear that the size of the epidemic would quickly overwhelm healthcare capacity.

5.19. On the morning of 4 March 2020, SPI-M-O member Riley called me. He explained what he believed would likely play out. Essentially, he thought that government would attempt to have the epidemic but would be forced to introduce generalised NPI when the scale of the problem became apparent – when "the first major London hospital falls over" were his words. As he was describing this scenario, I became convinced that this was a highly likely possibility, and that the country would end up in a very bad position. Although no formal work had been written, the main benefit of generalised NPI would be to prevent healthcare being overwhelmed so that to apply them using overwhelm as a trigger must be suboptimal. The delays between infection and hospitalisation (about 2 weeks) and between hospitalisation and death (about 2 weeks again) meant that interventions have to be introduced well before the epidemic has reached the point at which it is to be stopped. The pattern of exponential growth means that the numbers

at the point of optimal introduction are much smaller than intuition suggests (see Explainer below).

5.20. I urged Riley to bring a written paper to SPI-M-O. Riley emailed me his paper late on 7 March 2020 and shared with other members I believe [GM/119 - INQ000213049] [GM/120 - INQ000213050] I discussed further with Riley, and he forwarded the paper to SPI-M-O members and secretariat on 10 March 2020 [GM/121 - INQ000213052]. The paper was presented by Riley and discussed by the committee at the meeting 11 March 2020. I do not recall the details of the discussion, but it was clear that Riley was expressing an anxiety that most were feeling. Whilst Riley's paper was relatively early in development and was attempting a substantive task (to cross the epidemiology-economic interface) in the face of an epidemic, it has considerable merit in opening the discussion about the harms incurred by the virus and the measures to control it. In particular, it highlighted that not controlling the epidemic incurred considerable economic harms. I have no specific recollection of the paper from University of Edinburgh dated 11/03/20 ("Course of epidemic with 2 fixed period interventions…") being discussed at the same meeting, although we did follow Riley's paper with discussion of alternative more stringent strategies such as adaptive or repeated lockdown policies. The Edinburgh paper is a note illustrating the pattern of epidemic dynamics associated with multiple periods of more intense transmission reduction ('lockdown', LD). It demonstrates that multiple LD can reduce the peak prevalence, but prolongs the epidemic. The paper agrees with other work produced by SPI-M-O, and inspired more. It highlights the importance of timing and effectiveness of any such measures, as well as highlighting that there will be multiple peaks if the first peak is curtailed. Members of the SAGE secretariat and Cabinet Office attended that meeting as observers. However, there was no consensus statement produced from that meeting, and the Riley paper was not taken to SAGE. In my email exchange with Paul Allen I wrote that I did not wish the Riley paper to be lost and felt important that we raise the issues of strategy [GM/122 - INQ000213053].

5.21. Inspired by Riley, I also sent a set of notes that were included in the document amongst the SAGE 13 papers [GM/106 - INQ000106152] [GM/108 - INQ000213036]. This was not a SPI-M-O consensus (and never presented as such) but a "straw man" to stimulate discussion. My underlying concern was that there was too little thought being given to the strategy throughout the epidemic. I might have been wrong, and it was simply not being shared with myself and SPI-M-O, but the modelling community is the ideal group to inform such thinking.

Explainer: Exponential growth

5.22. Exponential growth arises in epidemic models because the incidence of infection (the number becoming infected per day) is a direct multiple of the prevalence (the number currently infected). Exponential growth is characterised by a constant doubling time. If we think of generations of infection (i.e., one person infected at generation one who infects people in generation two and so on) then the multiple is the reproduction number, which we were estimating as about 2. Consequently, the numbers of infection that we might expect in the first eleven generations of an epidemic are: 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024. These numbers get very big very quickly. The epidemic is made up of many such chains of infection, although what we observe is the total number.

5.23. This is the number of infections. The number of disease cases (e.g., symptom onset or hospitalisations) or deaths occur with a delay. For COVID-19 pre-vaccination, the delay between infection and death was about 4 weeks. If policy reacts to the numbers of deaths, and there is a one week doubling time, then the numbers of infections will have grown 16-fold.

5.24. There is another complication that because transmission dynamics and disease/death risk are heterogeneous then the infection process is not easily observed from the disease/deaths. This is essentially what happened in UK during the first wave; there was a large wave of infection largely among younger people in London with relatively little consequent disease, which then moved into older cohorts across the UK, including in care homes. ***[explainer ends]***

Section 6: April 2020 onwards

6.1. The quotes cited by the Inquiry ("we've kind of painted ourselves into a corner" and that lockdown was "a placeholder") are accurate and given during an interview with a journalist at the start of the first lockdown that appeared in The Times newspaper 4 April 2020. I went further in terms of discussing strategy than was prudent, but I was particularly frustrated at that point. The overarching reason was the lack of clear, long-term strategy for dealing with the epidemic. I also felt that there needed to be a more open public debate about where the harms of the epidemic would fall.

6.2. The main source of my frustration was the disconnect between political rhetoric and reality, as I saw it. Presenting LD as a solution to the epidemic did not seem to me to be optimal. It was a temporary measure to prevent the first wave overloading healthcare, and I was not sure that decision-makers understood that this would be a long process requiring management over many years.

6.3. It was also not clear to me how, having presented the LD as a solution, the UK would be able to reopen, and then re-introduce restrictions when the next wave started. SPI-M-O had extensive discussions about the use of NPI – the need, the potential impact, timing etc. – and it had become clear to me that restrictions should not be imposed unless it's clear how they will be relaxed, nor should they be relaxed without knowing how they would be reimposed.

6.4. I was also frustrated that complete LD had been introduced at one time, rather than gradually ramping up restrictions and learning what level of restriction would be required to control transmission. On the other hand, I was greatly relieved that LD worked. I was surprised, given the evidence that SPI-M-O produced, that schools were closed at that point.

6.5. However, I am not at all expert in how the population responds to restrictions of this nature, and I also learnt to distance myself from political rhetoric.

6.6. The issue of the role of LTCF ('social care facilities') and hospitals in maintaining transmission during April and May 2020 has also been addressed in paragraphs 3.97 to 3.109. My email to GCSA dated 17 April 2020 [GM/123 - INQ000213054] was inspired by data on registration of deaths in different settings. Although there were uncertainties in the interpretation of the data, I believed that they indicated a possibility that transmission was continuing in health and social-care settings and might be self-sustaining in those settings even though general community transmission was reducing. If this was the case, then the IFR in the sustained epidemic would be much higher because of the vulnerability of the population in health and social care. The major national control measure at the time was the general LD of the whole population, but if the majority of morbidity and mortality was being generated from transmission within and between hospitals and LTCF, then this would continue despite the LD.

6.7. This would be exactly the situation that policy should avoid, and I thought it right to alert GCSA. GCSA reassured me that this was understood and actively being

considered [additional exhibit – email RE_ OFFICIAL_ Hospital metrics [GM/123A] – **INQ000260625**].

6.8. This issue was specifically considered by SPI-M-O (20 April 2020) [GM/72 - INQ000213298] and included in the consensus that went to SAGE 27 (21 April 2020) [GM/45 - INQ000062295]. The consensus was, effectively, that it was a 'realistic possibility' that hospital transmission was self-sustaining. The consensus was that the data on deaths that I had seen, and that the situation in LTCF was sufficiently uncertain that no firm conclusions could be drawn. In the event, the transmission in these settings declined, so that either it was not self-sustaining, or measures were introduced to reduce transmission in these settings.

6.9. The other process that became clear in this and subsequent LD, is that chains of transmission continue after the introduction of LD. The role of LD is to prevent new chains starting. For example, in a household, transmission will continue during LD but will (ideally) not progress to other households. Similarly for hospital wards and care homes, LD does not prevent on-going transmission in those settings, but ideally prevents transmission into those settings.

6.10. I do not believe that I attended the small technical meeting on transport on 27 April 2020, although I was originally invited. In an email at 12:47 on 27 April 2020 I was asked by GCSA office not to attend [GM/124 - INQ000213055]. However, at 13:59 I sent an email to clarify as I had been sent an agenda by CO [GM/125 - INQ000213056]. I do not have a record of a response to the second email, and I do not have any recollection of the meeting. It remains in my diary as I was not removed from the invitation list. If I did attend then the people organising the meeting should have a record of my attendance, but I do not have any recollection or evidence that I attended. There were other meetings organised in the same series [GM/126 - INQ000213057], which I am certain I did not attend as I removed myself from the invitation list, i.e. I declined the meetings.

6.11. As already stated, I did not have a role in the decision-making, nor am I in a position to judge the correctness of the decisions. I can comment on the apparent role of scientific evidence in decision-making from my perspective.

6.12. The situation from May 2020 onwards was very different from that during the period leading up to the first LD. The principal difference was that it was clear that government did have the ability to be able to stem the epidemic, i.e., that transmission was

controllable. In some ways this made decision-making more difficult because it made the choice between transmission and measures to reduce transmission clearer. Had, for example, the first LD not reduced transmission to the extent it did, then there would have been no need for an explicit trade-off between transmission and measures to reduce transmission.

6.13. The outcome of the first LD removed a large uncertainty from the decision-making and gave more opportunity for development of strategies to manage the epidemic. Additionally, scientific understanding and public health data improved quickly. Specifically, there was a lot more knowledge about the risk factors for infection and disease. During the summer 2020 SPI-M-O developed the full range of modelling approaches (see paragraphs 3.47 to 3.68). By the end of August 2020, we had a varied range of tools to support decision-making.

6.14. A key role in decision-making is to ensure that decisions are evaluated, especially when similar decisions will have to be made in the future. Without evaluation sub-optimal decisions cannot be corrected, nor can data be accrued to support future decisions. This adaptive learning is essential. I give examples elsewhere in my evidence where opportunities were missed to enact policies that would have provided valuable data to improve decisions later in the epidemic, and, indeed, in future epidemics. My view is very much informed by the contrast in the decision-making process Feb-Dec 2020 compared to 2021 and beyond.

6.15. As discussed in Section 8, there could have been more productive use of modelling to provide evidence towards what the government wished to achieve. I had read in the media during summer 2020 (I cannot remember where or when) that January is, in economic terms, the least productive month and that therefore a LD was least harmful at that point. Had SPI-M-O been asked, for example, how to get through to January 2021 without a LD, or what was the minimum LD length to get through the winter without overwhelming hospital-care, I suspect that SPI-M-O would have been able to produce some meaningful evidence, and even the discussion would have been useful for both modellers and policy-makers.

6.16. My view of the government strategy document in May 2020 is that it is flawed and was a missed opportunity to explain more how policy intended to cope with the whole epidemic, rather than just the “unlocking” process. It gave the impression that the epidemic has happened rather than that we were 10-20% of the way through. My view

at the time was that there would need to be periodic LD of some form, but this possibility is not made explicit. At that point, without vaccines, the epidemic would have lasted may be three or more years. In the event several decisions were not consistent with this document. However, this is just my personal view, that an opportunity was missed to try and give an overview of the next years rather than describe the immediate future. For example, the inclusion of the RWCS would have given more insight.

6.17. In answer to the specific questions.

a. I have not read the book in question entitled *The Year the World Went Mad*, so do not know the context of the quotation. It is always possible, especially after the event, to say that more should have been done, and better. I can only comment on specific measures which might have been taken.

b. The 'Eat Out to Help Out' scheme was an example of SPI-M-O not being asked to provide information for some decisions. I do not recall it having been discussed on SAGE. This scheme was an apparent contradiction of the government aims laid out in the May 2020 strategy document of "The overriding priority remains to save lives" and acknowledgement that "the fewer contacts, the lower the risk". The timing of the incentives was just as the epidemic was starting to pick up, i.e., $R > 1$ and prevalence increasing. I fully understand that government had to balance between epidemiological and economic outcomes and that different sectors of the economy have different patterns and requirements, but from a transmission dynamic perspective, increasing the rate of transmission at that point was sub-optimal. I have not seen a full evaluation of the incentive, so it is not obvious what impact the incentives had on transmission, and it may be that government considered it to be 'cost-effective' in terms of increases in prevalence compared to the economic return.

c. Personally, the worst point of the epidemic was January 2021, when there were over 1000 people dying per day and the country was locked down. Objectively, given that COVID-19 illness and deaths and lockdowns are the two worst outcomes I believe it was the worst period. I am also of the opinion that it was avoidable. Whilst there was misfortune in terms of the timing of the appearance of the alpha variant, the increasing prevalence from July 2020 onwards led to having a second wave that was worse than the first.

6.18. I agree that the rationale for relaxation of measures after April 2020 appeared relatively random, and I was somewhat confused about government policy and intentions – I

was not sure what they would have regarded as a good outcome. My main concern at the time, and looking back, is that there was a missed opportunity for significant learning about transmission patterns. I understand that government would have found it difficult to implement different relaxation patterns in different places at different times, but this would have revealed a lot about the impact of different policies and policy-mixes that would have been hugely beneficial for the next two years, and especially in the autumn 2020 (see paragraphs 6.26 to 6.28).

6.19. Whilst decision-makers might have had a "belief" [as said by the Chancellor 5 November 2020] that TTI would keep $R < 1$ this was not clear and I, and others on SPI-M-O, thought that there would need to be some form of national intervention. Contingency planning is clearly important when outcomes are uncertain and different outcomes have very big implications. We were not asked, but I thought that the chances of TTI being able to keep R at about 1 were small. The best chance would have been if prevalence had been kept low – see paragraphs 6.29 to 6.34 and role of precautionary breaks in keeping prevalence low.

6.20. I am not sure what "staying ahead of the virus" means, although based on the May 2020 strategy document it meant keeping the reproduction number less than 1. It has been learnt in attempting to control other epidemics that 'top down' and technological solutions rarely work when the intervention requires people to modify their lives in order to reduce transmission. Technology can support people, as for example seen with the availability of testing in autumn 2021, but they learn how to use and adapt it rather than having it imposed. There were many ideas circulating in summer 2020, including mass testing and development of the NHS COVID-19 app, but neither of these would be a single answer. A commitment to evaluating different mixes of interventions would have both given invaluable information but also empowered the public to find their own solutions.

6.21. It is clear now that prevalence was increasing from end of July 2020 onwards (Section 9). The introduction of 'eat out to help out' at this time was inconsistent with the strategy of keeping the reproduction number below 1. By mid-September it was clear that the UK was only a handful of doubling times from a situation the same as the first wave and the evidence base for how to reduce transmission was little advanced from March 2020. The request to provide evidence for reintroduction on 17 September 2020 was with a very short turn-around but could have been developed over the summer with a more long-term, over-arching view of the epidemic (see paragraphs 6.26 to 6.28).

Explainer: Transmission interventions

6.22. There are essentially three ways of controlling transmission – vaccination, blanket reduction in contact (and thus transmission) and reduction in transmission targeted at people infected. Targeting measures at infections requires that infections are diagnosed, and different interventions such as quarantine and isolation can be used to ensure that those infections have a much-reduced onward transmission. Blanket measures essentially try to reduce the amount of relevant contact in the population but do not require diagnosis. Examples of blanket measures are campaigns to use condoms or wash hands. TTI is a targeted intervention and LD is a blanket intervention.

6.23. Targeted interventions depend critically on both the speed and accuracy of diagnostics, and the interaction with people. Implementing a programme requires both the technological function (i.e., personnel, laboratories etc) and an interaction with the public to ensure that the programme gets used appropriately. Both are essential. Barriers to using a diagnostic system and adherence to the rules/guidance are many, but stigmatisation and economic constraints are the most common. ***[explainer ends]***

6.24. My purpose in attending meetings to which I was invited was to provide expertise and to link with others in SPI-M-O, and to a lesser extent, across government. I did not organise meetings and did not keep minutes or notes of meetings other than reminders for immediate actions. Throughout the epidemic, the process of gathering evidence was where external experts were involved, and turning it into advice was largely done by government employees, especially through CMO and GCSA. The meetings the Inquiry has asked about were, to a large extent, to provide these individuals and their staff with supporting information and aiding interpretation.

a. The meeting 29 May 2020 was titled as a "general modelling catch-up". I cannot recall what was said that this meeting, although judging from the email from later that day summarising the actions, it would appear to have been about the interaction between SPI-M-O and the newly announced JBC [GM/127 - INQ000213058].

b. The meeting 31 July 2020 was a discussion on what areas SAGE might consider in the future. I do not have a record.

c. In the invitation the 7 September 2020 meeting was informal, and minutes were not taken. I have no recollection of what was discussed. However, I sent an email to GCSA after the meeting commenting that there is no formal way of judging success, which suggests that at least some of the discussion was retrospective [GM/128 - INQ000213059].

d. The 14 October 2020 meeting was a discussion of long-term scenarios, but I cannot recall any details.

e. I have 5 meetings in my diary 14 December 2020, all of which are about modelling, so I am not sure which is being referred to. I didn't organise any of them or keep any minutes.

f. This was an in-person meeting in Oxford 15 September 2021. The attendees discussed the experience of providing scientific advice during the epidemic. There was co-ordinated discussion around what had been learnt, i.e., it was retrospective. I do not have a record of the meeting.

g. I have no record of the 6 December 2021 meeting.

6.25. There was a very short turn-around time for consideration of the reintroduction of measures to control transmission (3 days), presumably because policy-makers required updated information on the potential interventions that could be used to reduce the exponential growth that was being seen with a doubling time of between 1 and 3 weeks (although a week later it was estimated between 9 and 14 days). There were two draft papers available on 18 September 2020 that I commented on [GM/129 - INQ000213061] [GM/130 - INQ000213062] (*Notes on NPI Batting Order* and *Draft list of national NPIs*). SAGE 58 was held 8:30am 21 September 2020 to discuss the two papers. In the event, I am not sure that I contributed more than other members of SAGE to the papers.

6.26. An effective summary is in the SAGE 58 minutes [GM/60 - INQ000212102]. There was much uncertainty in the impact of different interventions on *R*, which was the metric used to measure the impact on transmission. Some of the uncertainty could have been removed with planned imposition and relaxation of interventions around the first wave – one of the failures was to learn more about how to control transmission during this, and future, epidemics. Paragraph 12 in SAGE 58 minutes makes this explicit.

6.27. I had considerable anxiety at the time that the UK would find itself in the same position as March 2020 with a high prevalence of infection, morbidity and mortality and be forced into a damaging, national LD. Preventing the prevalence from getting to high levels would prevent the need for this action. Much to my surprise and consternation, there was considerable public denial of the possibility of another large wave requiring LD to control.

6.28. I believe that the concepts of a 'circuit brake' or 'circuit break' and a 'precautionary break/brake' are subtly different. In electronics, a circuit breaker 'trips' when activity passes some threshold to prevent overload. The idea of a precautionary break is that it applied regardless of the epidemiological situation, but timed to cause the least economic and societal harm. The idea arose during SPI-M-O discussions over summer 2020 about the optimal timing of lockdowns. The start of waves are exponential processes (see Explainer at paragraphs 5.22 to 5.24), which LD stops. The first LD had been introduced when it was unclear what impact it would have on transmission, and we learnt that it reversed exponential growth (into exponential decline). In effect LD acted to reduce prevalence, thus setting the epidemic back to an earlier time in the exponential growth.

6.29. The point of precautionary breaks is that they are implemented before restrictions are necessary, i.e., when the prevalence is relatively low to prevent it getting high. The aim is to keep prevalence low whilst it is low, in contrast to emergency breaks which aim to reduce prevalence once it becomes high. I had read somewhere in the media that a significant negative impact of LD was not knowing when they would be introduced, nor how long they would continue as well as what the actual interventions would be. The uncertainty of 'when', 'how long' and 'what' impacts on both economics and individual well-being. If LD are introduced as an 'emergency break' to prevent crashing into a disastrous situation, then they are not pre-announced and they are of indeterminate length. Theoretically, LD have the same effect on transmission if they are applied early or late and applying them early means that they can be pre-announced and for a fixed period of time. In effect, the theory is that a precautionary break acts in the same way on transmission as an emergency break, but with less negative impact.

6.30. There is uncertainty about how the population might react to LD introduced when prevalence, morbidity and mortality are relatively low. In particular, the aim of a LD is to reduce contact, and there is a risk that contacts will be deferred rather than stopped

if the dates of a LD are pre-announced. There was also uncertainty about the impact of individual restrictions on transmission (see paragraphs 6.26 to 6.28). Consequently, it seemed to me, to be prudent, if there was to be a precautionary break, to ensure it had maximum impact, and that suggested that doing it over the autumn half-term, when schools are closed anyway, would be optimal. If half-term were extended for a week (so two weeks in total), then it might be possible to have a LD of the same impact on transmission as the first LD but with reduced economic and well-being impact.

6.31. If this was to be done, and the pre-announcement was to be worthwhile, then it would have to be announced well before the half-term (which was end of October in England). School closure dates vary in England, but the great majority occurred in either week 43 or 44. School closures in devolved administrations vary considerably more, so that this suggestion was largely for England alone.

6.32. The term 'circuit breaker' was coined as meaning a threshold to trigger interventions as early as May 2020 [GM/131 - INQ000213064] The implementation SPI-M-O discussed was different, in that it would be applied regardless of the prevalence, but rather in response to an increasing epidemic. I cannot recall the exact date on which the idea emerged, but was agreed as a possible strategy during the SPI-M-O meeting 16 September 2020. The idea was developed in papers presented to SAGE 57 and 59 and the SAGE 59 SPI-M-O consensus statement [GM/132 - INQ000213065] [GM/133 - INQ000223535] [GM/51 - INQ000120558] [GM/134 - **INQ000215660**] [GM/61 - INQ000213284]. The work was made public as a preprint 14 October 2020, and subsequently published [GM/135 - INQ000213068] [GM/136 - INQ000213069].

6.33. I am not sure what Professor Woolhouse means as "alternatives to lockdown" since the term LD is not clearly defined. The papers to SAGE 58 discuss the specific interventions that could have been introduced in various mixes. The precautionary break was proposed as a limited time 'lockdown' with the suggestion that it could have a much higher positive/negative impact ratio [GM/60 - INQ000212102].

6.34. By the 14 October 2020, with half-term starting 19 October in many parts of England, any value in pre-announcing a precautionary break had been lost, in my opinion, hence my comment that "we had missed the boat" for a precautionary break. My comments came from a press conference to science correspondents organised by the SMC following the publication of the SAGE 58 papers [GM/60 - INQ000212102] (see paragraphs 6.26 to 6.28).

6.35. My comments "to build a national narrative about" the epidemic were based on my wish that there had been a public discussion and agreed strategy for getting through the whole epidemic rather than dealing with the epidemic one week at a time. A precautionary break would, to my mind, have been a means of gaining control over the epidemic, rather than having the epidemic dictating emergency measures. Had a precautionary break been successful then we may have been able to avoid the situation in January 2021 when we had both very high morbidity and mortality and LD (see paragraphs 6.11 to 6.17). Potentially, as a country, we could have navigated a path of much lower morbidity and mortality with less LD disruption. However, it is also possible that the precautionary break would not have worked in the same way as the emergency LD that were employed.

6.36. In February 2020 I had asked on SAGE what level of spatial granularity would be relevant to policy so knew that it would likely be a feature of response. Even if each nation of the United Kingdom operated the same policy across its jurisdiction it would create variation within the epidemiological unit of the United Kingdom. Pre-pandemic, SPI-M had briefly considered the possibility that devolved nations could be caught in theoretical gaming, where, for example, the best policy for Scotland depended on the policy in England and *vice versa*. Devolving strategy decisions is often sub-optimal unless there is a firmly agreed common goal.

6.37. SPI-M-O looked at regional variation expected during an epidemic in February and March 2020. By June we were analysing data at a relative fine spatial scale (e.g. SAGE 40 4 June 2020 paper) [GM/137 - INQ000120526] [GM/138 - INQ000213072]. Although national restrictions were relaxed over the summer, there were various restrictions locally, especially in Leicester. SPI-M-O was estimating the epidemic pattern regionally. SPI-M-O was asked to consider the differential impact of regional restrictions in July/August 2020 (SAGE 50 6 August 2020 paper) [GM/139 - INQ000213073] [GM/140 - INQ000213074]. I had some discussion with Alan Penn, CSA Ministry of Housing, Communities and Local Government (MHCLG) about modifying restrictions to localities during the first LD, but they had not come to SPI-M-O.

6.38. Tiers were introduced 14 October 2020 without any further discussion on SPI-M-O, and I do not remember any on SAGE. Observing the patterns in localities, such as Leicester, where additional restrictions had been introduced did not give me huge confidence that they would be successful at bringing *R* below one. With some benefit

of hindsight, I suspect that this is because the spatial scale, and in particular the 'edge effects' given that the localities were not separated. There was spill over from one area to the next, so that high incidence in one place resulted in higher incidence in connected (e.g., neighbouring) places. SPI-M-O did not have access to good quality mobility data (e.g., from mobile phone companies) so could not say what the connectance between areas was, or what an optimum spatial granularity might be. Because an area had to have high incidence before additional measures were introduced, this meant that localities tended to move up the tiers. With hindsight, although there was some discussion on SPI-M-O at the time, a better approach might have been to relax measures in places that had low incidence. This would mean making the default Tier 4 and moving down the tiers if incidence fell. Tiers as implemented certainly had an effect (see SAGE 69 19 November 2020 paper) [GM/141 - INQ000120569] [GM/142 - INQ000114473], and, we now know, might have been sufficient had they been introduced earlier. Nonetheless, at the time, the prevalence of infection had been increasing for several months and I was, at the time of the interview, pessimistic that the tiers system would be sufficient to regain control.

6.39. The Inquiry has asked about a particular email from SPI-M-O co-chairs. The email to CMO and CGSA 23 October 2020 came about after a discussion that co-chairs and secretariat had on the evening of 22 October [additional exhibit – email "How long have we got" [GM/142A – INQ000260626]. The context was exponentially increasing prevalence – not completely consistent regionally we now know because of the impact of tiers. The purpose of the email was to indicate to policymakers that time was getting short if they wished to avoid a larger wave of hospitalisation and death than had occurred in the first wave. In the email conversation that followed, CMO indicated that this would be more useful regionally. However, I am not sure that SPI-M-O did this due to pressure of time. No further response was received by me.

6.40. The question of whether the government was insufficiently cautious depends entirely on what the strategy was. It was a commonly held view on SPI-M-O that everything worked better at a lower prevalence, especially control of nosocomial transmission (and transmission in other settings controlled by testing) and contact tracing. As the prevalence increased, so the UK lost control, and had to resort to LD. I do not know if this was intended, although as in paragraphs 6.29 to 6.34, my view at the time was that shorter, pre-announced LD to keep prevalence low would have been less damaging.

6.41. Overall, the autumn of 2020 stood in stark contrast to January 2021 onwards (see paragraph 8.43 to section 9 for more detail). During that period SPI-M-O were told when decisions would be made and the evidence required for those decisions. Consequently SPI-M-O were able to say what data would be required to create the evidence. In the case of the tiers policy, we were not informed that they would be applied and did not have the necessary data to evaluate them until several weeks later (see the data comments added to the tiers evaluation paper on page 1) [GM/142 - INQ000114473].

6.42. In an interview that I gave to the BBC on 13 July 2021, six days before all major restrictions were to be lifted on 19 July 2021, I said that daily hospitalisations were likely to reach between 1000 and 2000. This view arose from consideration of the modelling from SPI-M-O. Vaccination had been widely rolled out, so that the wave would cause less morbidity and mortality than previous waves. Vaccine protection against infection (and onward transmission) was relatively low, and the appearance of the delta variant had increased the chances that a wave of infections could be large. The SPI-M-O modelling summary presented to SAGE 93 7 July 2021 was the primary source [GM/143 - INQ000213077] [GM/144 - INQ000120629].

6.43. I was also aware that the models SPI-M-O were using tended to be more 'peaky' than the observation. The estimates of the total size of a wave are quite good but forecasting peaks is inherently difficult. I said in the same interview "…even if we don't get up to very high numbers, the numbers that we get up to might last for several weeks, six weeks or so, in which case there's still a considerable burden on healthcare." In the event, the wave of infections turned into a long but surprisingly stable period of about 800 hospital admissions per day in England.

6.44. None of the models captured this long period of stability. It is not known what caused it, although there is circumstantial evidence suggesting that it was people's modification of behaviour based on the information gathered through testing and the use of the NHS COVID-19 app. The period immediately following 19 July 2021 (the 'pingdemic') had the same reduction in transmission that had been seen during lockdowns as presented in SPI-M-O consensus at SAGE 96 14 October 2021 [GM/145 - INQ000213079] [GM/146 - INQ000120650]. As stated previously, models are unable to predict the precise future trajectory as human behaviour is unpredictable.

Section 7: Access to, Sharing and Quality of Data

7.1. The Inquiry asks about data provision for modelling. I have given some explanation of why data are required and the typical issues surrounding it in related Explainers below. The Inquiry should also note that my position meant that I was not using the data (nor supervising people using the data), so I was not immediately aware of when data were available. I did not own any of the data, nor did I have the resources to make it available. My role in terms of data was to alert government (primarily the SPI-M-O and SAGE secretariats) to problems. I did not keep a record of which data were available in which format at which time, but I believe that the SPI-M-O secretariat did.

7.2. In February 2020, the primary concern was to have access to as detailed data as possible from the initial cases, which were in China and more closely linked administrations and nations such as Hong Kong and Singapore. The minutes from SAGE 4 February 2020 refers to accessing data from these settings [GM/113 - INQ000051925]. In particular, estimates of the biological delays between infection and symptom onset, symptom onset to hospitalisation, and hospitalisation to death/discharge.

i. The requests for these data were co-ordinated by the deputy CMO and FCO. SPI-M-O were asked to provide a prioritised list of the data, as summarised in the email thread 7 February 2020 [GM/147- INQ000213081] . Additionally, the outbreak on the cruise ship *Diamond Princess* was an important source and the SPI-M-O co-ordinated interaction with the Japanese public health authorities as summarised in this email thread [GM/148 - INQ000213083]. At the same time, SPI-M-O members were also using data on flight volumes and arrivals into UK from China that the SPI-M-O secretariat had procured. The Department of Education also shared data from 17 February 2020 [GM/149 - INQ000213084].

ii. Given the situation, the delays in data provision were understandable, if frustrating at the time. If the data had been provided more quickly then the results would have been generated more quickly. I am not convinced that this would have changed outcomes significantly, given that the UK was a few weeks behind the epidemic in other countries. There was a global effort to gather and analyse data during February 2020 which meant that UK was not relying solely on the results obtained from analyses within the UK.

iii. The reasonable worst-case scenario (RWCS) is a scenario such that most outcomes in reality fall below it, i.e., if organisations plan for the RWCS then it will most likely have planned for the outcome. During February 2020, given that the information and data was incomplete, SPI-M-O recommended that government continued to use the pre-pandemic RWCS from the SPI-M modelling summary. As data and analyses accrued, SPI-M-O adapted this generic RWCS to be COVID-19 specific, for example the SAGE 6 10 February 2020 paper [GM/150 - INQ000213085]. By 3 March 2020 the RWCS for COVID-19 was in draft format and was endorsed by SAGE 14 10 March 2020 [GM/151 - INQ000109125]. I do not believe that this was held up significantly by data availability from outside the UK.

iv. I do not have the expertise or experience to comment on non-UK organisations' data streams. If the question relates to the UK, then please see the following section at 7.6.

v. The Inquiry asks specifically about the lack of information on the duration of immunity and its impact on accuracy of modelling. At the time (February 2020) there was no data available as time since first infection was at most only 2 months, so duration of immunity had to be judged based on virology and other coronaviruses. SPI-M-O models initially assumed that immunity is full and lifelong (resulting in a so-called SIR modelling framework) which would have been accurate for the first few months of an epidemic in any situation as the vast majority of infections would have been primary infections. NERVTAG produced a consensus view of immunity 27 April 2020 [GM/152 - INQ000213087] (document: NERVTAG: View on SARS-CoV-2 protective immunity) and SPI-M-O members were exploring the potential (and at that point theoretical) impact of waning immunity during the second half of 2020. In the event, this did not start to become important until summer 2021, although SPI-M-O members were just as interested in the duration and protection from vaccine-acquired immunity as immunity derived from exposure and infection. I do not believe that any more could have been done to address the question of duration of immunity, and I do not think that it materially altered the scientific evidence generated. However, if government had asked SPI-M-O to consider the whole epidemic at the start and produce scenarios over the coming 5 years (see paragraphs 8.43 to 8.67) then we would have had to include different

assumptions about immunity duration. There are many other processes for which data are not possible during an epidemic of a novel pathogen, and this will be the same during the next epidemic. For example, if immunity is not complete and life-long, what is the IFR of repeat infections? What proportion of people, if any, can become long-term persistently infected? This is not an unusual situation, even for endemic infections that have been studied for many decades. Models are well placed to understand the potential impact of different processes given that they make assumptions explicit and can be used to highlight those data that are required to solve the most influential processes.

vi. I was not using the data myself, so cannot comment on how well described it was, nor how interactive the data owners/providers were. I do not remember hearing any significant problems, and do not have any emails relating to these data.

Explainer: Models and Data

7.3. Models are theoretical constructs that consist of a structure and parameters. The structure is based on mechanistic understanding and is usually based on a mixture of assumptions and observations. The parameters allow the model to be fitted or tuned to particular circumstances. For example, we know that people contact differently by age-group from study of other diseases, so this can be included in the model structure, but then has to be fitted to the COVID-19 data as contacts were altered by restrictions and guidance. Without data, models are theoretical 'toys' – you can learn a lot from them, but they are not guides to specific policy development.

7.4. As models are infinitely flexible, they can be fitted to most data. The fundamental is that the data must be well described so that it is clear which part of the model they are related to, or the model can be adapted to fit to them. The clearer to the modeller how the data are collected and what they mean, the less likely that there is a mismatch between data and model.

7.5. There are substantial legal frameworks surrounding data usage. Data are owned and are shared within legal contracts. Typically accessing data also requires ethical approval in a research context. This normally takes considerable time and effort to arrange. Modellers very rarely have data collected specifically for modelling but are usually using data that is collected for other purposes. ***[explainer ends]***

7.6. The SPI-M pre-pandemic modelling summary (annex 4) emphasises the data required for an epidemic as being two principal types: individual (line-list) data and aggregate data (i.e. anonymous, unlinked time series of numbers) [GM/14 - INQ000212124]. The individual data were to be generated by the FF100 system, which tracks the first few 100 cases and their contacts. This system worked fairly well, although was clearly under-resourced, with a consequence that the completeness and timeliness were severely compromised. To the best of my knowledge, the system for aggregate data had not been developed pre-pandemic, so had to be created *de novo*. Whilst the data streams were set up, a major issue was the delivery of the data to SPI-M-O members, which is largely the subject of Section 7. Here I give a history of events relating to early data.

7.7. SPI-M-O members received information on data sharing from PHE 17 February 2020 [GM/153 - INQ000213088]. Although a much-streamlined process compared to normal access to PHE data, this nonetheless required that Universities agreed contracts. I know in one case that this took several weeks, hampering the development of forecasts. It should also be noted that this outline only applies to England. I had raised the question of how SPI-M-O would operate between and across the four nations pre-pandemic, and it continuously raised issues throughout the epidemic.

7.8. The initial plan was for PHE to co-ordinate and deliver health-related data, as described in this email [GM/154 - INQ000213089]. I received the first FF100 data set 6 March 2020 and daily after that, usually emailed late in the day and accessed via the egress system. There were considerable problems with the data. In particular, the delay to adding cases to the database seemed to be about 2 days. Additionally, much of the information, and especially onset date, was missing, and several data sources were conflated. The length of reporting delay was especially concerning as it made the nowcast estimates of current incidence very uncertain, as highlighted to SPI-M-O secretariat by email early 12 March 2020 [GM/155 - INQ000213090]. After a meeting with SPI-M-O secretariat, I emailed PHE 12 March 2020 highlighting these concerns [GM/156 - INQ000213091].

7.9. There was considerable confusion about data definitions, especially how to interpret 'sporadic', highlighted in an email to GCSA 15 March 2020. The number of 'sporadic' cases was growing rapidly. The definitions resolved quite quickly – as indicated in this email thread of 15 March 2020 [GM/77 - INQ000213304], and email alerting CMO and GCSA 16 March 2020 [GM/157 - INQ000213092]. The FF100 system was failing as a

surveillance system, although it was not designed for that purpose. Given the rapidity of the development of the epidemic, and the consequent hospitalisations, and the lack of resources for community testing and contact tracing, and the lack of serological testing, hospital data was the most obvious data switch to.

7.10. An email thread 16 March 2020 initiated the provision of data from the NHSE dashboard data, which initially came through the CHESS system [GM/158 - INQ000213094]. My first receipt was 16 March 2020. I highlighted data concerns to deputy CMO on 17 March 2020 email [GM/159 - INQ000213095]. By 19 March 2020 SPI-M-O members were producing forecasts of ICU beds that would be occupied in the coming weeks, and the SPI-M-O consensus statement of 20 March 2020 (SAGE 18 23 March 2020) highlights that the epidemic was growing rapidly [GM/79 - INQ000213306] [GM/160 - INQ000052717].

7.11. I emailed deputy CMO 23 March 2020 forwarding him a paper from the Lancaster SPI-M-O modelling group which laid out the data requirements and problems, particularly in terms of spatial patterns [GM/161 - INQ000213097] [GM/162] - INQ000213098.] There had been an implicit change of strategy from avoiding an epidemic in autumn and winter to preventing healthcare capacity being exceeded. In order to do the latter, we needed to ensure that the epidemics in different places did not exceed capacity placing additional emphasis on spatial patterns – see my email 19 March 2020 to SAGE secretariat [GM/163 - INQ000213100].

7.12. Given the confusion over the data streams, data on deaths became a more valuable foundation. These data are collected routinely, so that the systems are robust and well-understood. However, the notification and reporting delays are long. Nonetheless there were issues with deaths reported via PHE (see email 28 03 2020 [GM/164 - INQ000213101]), and there was concern that there might be double reporting, i.e. the same death reported multiple times. This was raised with ONS, and they were very helpful in ensuring that SPI-M-O members had accurate, timely data.

7.13. I emailed CSA for DoD about the state of the data 25 March 2020 as there were still substantial problems [GM/165 - INQ000213102]. LSHTM and Imperial College had, to their credit, got access to the data through a different route, but this meant that SPI-M-O was struggling to function as a group of modellers generating an ensemble. This gave rise to some frustration from other modelling groups. However, there was a suggestion that the NHS data could be made available through NHSX – see email

discussion 25 March 2020 [GM/166 - INQ000213103] [GM/167 - INQ000213104]. The NHSE sitrep was emailed to SPI-M-O members 26 March 2020 almost daily until May 2020 when they were on a secure server within DHSC that members could download. Scotland, Wales and Northern Ireland were added during April 2020.

7.14. Although there was some progress, it was significantly slower than policy was evolving. I emailed SPI-M-O members 29 March 2020 to get an overall view – the document referenced is one of the email responses highlighting some of the data issues. The email also highlights the value of the CO-CIN data [GM/168 - INQ000213108].

7.15. In terms of the CO-CIN data, I emailed the deputy CMO (and at the time co-chair of SPI-M-O) 14 March 2020 about access to the data with the observation that this could be a valuable data stream that could repurpose as surveillance [GM/169 - INQ000213109]. SPI-M-O were emailed about access 24 March 2020 [GM/170 - INQ000213110]. CO-CIN was a research study rather than management data from NHS that modellers were using as a surveillance stream. The CO-CIN data was particularly important for understanding the role of nosocomial infections (see paragraphs 3.97 to 3.109).

7.16. The existence of different data streams and systems for the four nations in the UK was a problem for the setting up of data delivery in March and April 2020 and continued to be an issue throughout the epidemic. Several times in the epidemic, for example, the administrations in the four nations used slightly different definitions of data, for example 'bed occupancy', which meant that the data being provided had different interpretations. Different regulations and different academic year dates also had to be taken into account. SPI-M-O kept a full list of different measures in place at different times under different administrations, but I cannot locate it.

7.17. The healthcare (hospital) data were essential to modelling the epidemic, given that the strategy became not to allow the epidemic to go above capacity. The CO-CIN data was important for our understanding of nosocomial infections, but not essential to surveillance; its particular merit was for understanding risks and treatments in contribution to medical understanding. The PHE data changed hugely as a result of the availability of testing, and this became a central data stream. Combined with the viral sequencing data (and the opportune switching of S-gene positivity) this gave unprecedented data on the impact and spread of the three major variants that arose (alpha, delta and omicron). The PHE data streams also proved flexible so that as the

epidemic progressed further data were added, including all positive tests (not just the first), all negative tests and viral sequence data.

7.18. The ONS CIS study, in parallel with the REACT study, were also essential data streams to guarantee the testing data. SPI-M-O did not get any data on the 'reason for testing' despite many requests. The testing system was designed as an intervention rather than a surveillance data stream, i.e., testing was available to modify behaviour and reduce transmission, so that we were not sure it would be providing unbiased data. The randomised community studies provided a different view of the epidemic.

7.19. In terms of data that were not provided, there are three major sources of data for which SPI-M-O asked, but which did not materialise. First is mobility data, which measures the rate of movement between places and therefore is an important piece of information for spatial transmission. There were data sources from Facebook and google and testing data, but not at a very fine scale. This email thread of 18 November 2020 gives an indication of what was available and some shortcomings [GM/171 - INQ000213111]. Essentially, SPI-M-O members estimated the interaction between places through the epidemic, which is a lagged indicator. It is possible that having timely high-resolution movement data would have improved the ability of the models to forecast the short-term (i.e., MTP) but would not have improved the scenario modelling.

7.20. The second data that we asked for, and went to several meetings about, was data from the TTI system. In the 19 August 2020 SPI-M-O consensus statement, there is a request for "detailed test and trace data" [GM/172 - INQ000213112]. It was very important that TTI worked as well as it possibly could, and SPI-M-O members were very keen to help. We discussed how TTI could develop its own models but were not invited to support closely and did not have access to the data.

7.21. The third data that would have been very informative, as in the 19 August 2020 SPI-M-O consensus statement, is the inclusion of additional fields in the pillar 1 and 2 testing data: "… would be substantially more informative if they included the reason the subject had come for a test – this is critical information for interpretation of the incidence of positive tests" [GM/172 - INQ000213112].

Explainer: Data Formats

7.22. The ideal data consists of line lists in which each individual person's complete data are recorded together. Typically, this will be some individual identifiers (age, sex, place etc) and a series of dates (test results, hospitalisation, death etc). In the case of outbreak data, there might also be information on how individuals are related to each other. However, such data are very unusual and pose considerable access problems given the identifiability and lack of anonymity. Without line lists, data typically consist of time series of numbers, for example the numbers of people who tested positive by day and the numbers of people admitted to hospital by day. Some of the people in the two data streams will be represented twice (test then admission) but we cannot tell who is who.

7.23. Line lists can be recreated by being able to link across datasets. For example, if the testing and admission data streams record individual identifiers (name, date of birth etc) then it is potentially possible to look back and see when an admission tested positive. ***[explainer ends]***

7.24. On 18 March 2020 NHSX were proposed as the platform that would be able to provide consistent data streams. I was asked to provide comments on what was required and responded [GM/173 - INQ000213113]. My email 24 March 2020 shows that there was no solution although it was still in the hands of NHSX [GM/174 - INQ000213114]. At a SPI-M-O subgroup on 27 March 2020 there was some discussion as indicated in the SPI-M-O secretariat note of the meeting [GM/175 - INQ000213115]. Dstl were already providing support for SPI-M-O to combine the short-term forecasts and were becoming involved in data redistribution.

7.25. It was not clear to me that the data situation was being communicated to the people (within government) who could solve the problems. Given the pressure that SPI-M-O was under in terms of understanding the epidemic and providing information that was being relied on for decision-making I emailed a summary of the situation to GCSA 30 03 2020 [GM/176 - INQ000213116] [GM/177 - INQ000213117]. There were many emails, but, from my viewpoint, little progress [GM/178 - INQ000213120]. The email from DoD was very welcome news [GM/179 - INQ000213121]. There was then a flurry of activity (emails on my part), and it appeared that a sticking point was legal access to the data and GDPR requirements as indicated in this email [GM/180 - INQ000213122]. The Dstl gained access to the NHSX data 10 April 2020 as notified

[GM/181 - INQ000213123], By 22 April 2020 the Dstl system was active and SPI-M-O groups were gaining access as in this email [GM/182 - INQ000213124] [GM/183 - INQ000213125]. The Dstl system continued until January 2022 when UKHSA took over the process.

7.26. Overall, this period felt very 'operational' rather than simply providing scientific evidence. The critical things we needed to understand in early March 2020 were how big the epidemic was and how fast it was growing, and the data available did not allow those to be estimated. It is clear to me now that there was a gap in the surveillance planning which emerged partly because of problems with serology data, and partly because there had been no prior consideration of a policy option of stopping the epidemic. The serology for an influenza pandemic was expected to be much more informative, allowing it to take over when FF100 stopped, whereas that for SARS-CoV-2 is far less informative at both an individual and a population level. It took about 6 weeks to go from 'no system', to delivery real-time information to SPI-M-O members, which was a remarkable effort by all those concerned.

7.27. With hindsight, a lot of the confusion in early March 2020 was caused by the huge influx of infections from France and Spain, as shown by genomic analyses [GM/100 - INQ000213028]. The 'mental model' that I had, was of a number of small seeding events arising from travellers from countries that were recognised as having a higher prevalence. The surveillance system was not set up to detect a large influx, and it confused the data greatly, and I suspect, created much of the confusion during March 2020. However, I do not think that this failure of data made material difference to policy decisions as the effect of a large influx was the same as on-going transmission.

7.28. My comment that "Reporting delays play havoc with data streams and make them very difficult to analyse in real time. If the delays change or vary by group, then they can distort analysis a lot. My comment "I wonder what these will do to the *R* estimates next week" on 5 October 2020 was in response to an apparent problem with reporting cases that had been noticed by an analyst on twitter [GM/184 - INQ000213127]. I was making the general point that if the reporting delays in any continuous process change, and the change is not separately observable, then the data can easily be misinterpreted. Reporting delays are inevitable in epidemics, but if they remain constant, then they can be accounted for.

7.29. As a specific numerical example, if we know (from experience) that 50% of hospitalisations are reported within one day, then we can compensate for that delay by doubling the number of hospitalisations reported. But if the reporting delay lengthens, so that only 25% are reported on a given day, then the change in reporting delay will be interpreted as a drop in hospitalisations. Eventually the change in delay will be seen, but during a real-time understanding of the data, such changes can create considerable confusion.

Explainer: Data Delays

7.30. Delays in data are inevitable problems in real-time modelling. Some delays are biological (e.g., the times between infection, hospitalisation and death which can alter with changes in clinical management), and some are administrative (often termed reporting delays). Biological delays can be characterised, understood and measured through line list data and linkage between data sets (see below). Reporting delays can be characterised if the date of observation is recorded. A key issue for reporting delays is that if they change then they cause significant problems because there is a delay in knowing about the delay. Different data streams have different reporting delays, as well as different sensitivities in different groups, so that separation of streams (i.e., knowing how and why cases are identified) is important. **[*explainer ends*]**

Section 8: The Use of Modelling during the Covid-19 Pandemic

8.1. Individual models are abstractions of reality and not 'truth'. Individual models vary in terms of the assumptions (both implicit and explicit), the data used and their coding. The quote from George Box that "all models are wrong, but some are useful" is often used, largely because it is correct. Models are abstractions of reality. Models are only worthwhile if they are simpler than reality, i.e., they capture important dynamics and processes that give insight and understanding to what is being modelled. Models can also be built in order to predict (see paragraphs 3.47 to 3.68), but again, inclusion of more detail and data is no guarantee of increased predictive power.

8.2. Models of infectious disease transmission dynamics are very different to models used in engineering and physics. Infectious disease, and indeed any biological system, is inherently non-linear; for example, people's behaviour is not fixed, but is dependent on other's behaviour and on what has happened. Popular perception of models is more

akin to models in engineering in which parameters can be measured accurately and models developed to predict the load a bridge will bear, or whether the aircraft design will fly. However, models of infectious disease are a long way from this.

8.3. Consequently, no single model in infectious disease is 'correct' in the sense that a drawing of an elephant is not correct. If somebody wishes to know what an elephant looks like, then they will generally have a better idea if they consider multiple drawings of an elephant by multiple artists, rather than just relying on one view. This begs the question of 'how good does a model have to be before it can be considered sufficiently correct?'. I do not believe that there is a definitive answer to this question. Modelling building is an 'art' – which structural variables to include depends on what is likely to be important both in capturing the data and in answering policy questions. Models can be formally fitted to data, and be shown to capture past dynamics, but there is no threshold above which a model can be labelled 'correct'.

8.4. During the epidemic, given the urgency of modelling, it was possible that models had errors and modellers had different approaches and used different data, so relying on a single model was not sensible. Having multiple models to address the same question provides more robust evidence. If the models agree on the important aspects, then there is more confidence in the evidence that the different abstractions (models) are providing. If the models disagree, then there is an opportunity to understand the role of different assumptions and data.

8.5. SPI-M-O did formally combine its estimates of the reproduction number, short-term forecasts, MTP and MTPS (see paragraphs 3.53 – 3.63). The combination results in a single projection, with the variability created by both the variability within each projection and between projections. It was agreed that each model or estimate would be given equal weight. There was some discussion during the epidemic about weighting models on the basis of past performance (i.e., the accuracy of past projections) but this was not agreed for good reasons. It would be useful research to go back to these ensembles and see whether a different combination method would have been better.

8.6. In terms of scenario modelling, multiple models are increasingly used in decision-making, in infectious disease and other spheres such as global climate change [GM/185 - INQ000213128] [GM/186 - INQ000213129] [GM/187 - INQ000213131]. There are many ways of making comparisons between multiple models, varying from

technical, statistical comparison to simple inspection of different patterns. Pre-epidemic the SPI-M approach had been to have a 'champion' model that was to be used as the pre-eminent source of evidence, and a number of 'challenger' models to, effectively, guarantee the robustness of the champion model. I believe that this approach is used in terms of evaluation of decisions for vaccination on JCVI. The alternative approach is to have a group (ensemble) of models of equal standing, with none having *a priori* pre-eminence. The primary rationale for choosing the champion approach is that the champion model can be resourced so that it is guaranteed to run. Pre-pandemic the champion model was to be the PHE model.

8.7. Early in the epidemic, it was clear to me that resource constraints in PHE meant that the champion model was not geared up to act as such. The immediate participation of the other modelling groups meant that, by default, SPI-M-O would be using an ensemble approach. I am not aware of any formal work on the optimum number of models in an ensemble, but given the concept of 'triple modular redundancy' in computing, I set a target of a minimum of three models providing evidence, although rather pragmatically, and to particularly avoid the situation of having two models that disagreed not knowing why or which one to believe. Given that SPI-M had not planned for SPI-M-O to conduct ensemble modelling there was no prior work on combining models. It was discussed early March 2020 [GM/188 - INQ000213132], but there was not the time and resource to set it up for scenarios modelling. Consequently, SPI-M-O scenarios were compared informally. With hindsight I think that this worked well, in that the consensus statements were largely text written for non-specialists, which would have been more difficult if there had been another layer of results between the models and the policymakers.

8.8. There was some thought early in the epidemic of making a 'super model' to include epidemiological and economic processes, and run multiple models. I and other attendees of SAGE thought this was going to be problematic to set up in April 2020 – see email [GM/189 - INQ000213133]. It is possible that such a framework would be useful, in which case the time to set it up and build the infrastructure is between epidemics, i.e., now.

8.9. The consequence of the ensemble approach that SPI-M-O adopted can be seen in the consensus statements. The *SPI-M-O: Consensus view on the potential relaxing of social distancing measures, 4 May 2020* shows four scenarios produced by different groups up to 1 October 2020 [GM/33 - INQ000213253]. None of these is a prediction,

but a clear sense of the likely pattern of the epidemic over the next five months can be seen, and they are qualitatively accurate in terms of what happened. The consensus / ensemble approach was particularly used in the period after January 2021. The aim is to extract the policy-relevant information from the combination of the models. Note that the usefulness of a model or ensemble is not determined by the model, but by the user or policymaker. Whilst modellers can try and produce evidence to support and inform particular policy choices, the evidence is used by the decision-makers not the modellers. Modelling works best when there is a clear question, i.e., the policymakers have set out the parameters and constraints of their decision.

8.10. Modellers understand the limitations of their models best, so that the proper use and interpretation of models and ensembles requires discussion with the user to ensure that the models are appropriate for the use to which they are being put. It is commonly understood in the scientific community that the ideal way of developing evidence from modelling is through a 'modelling cycle' within which the output from models is considered within the context in which they are to be used, which inspires development and refinement of models to produce more output etc. Time and resource constraints and the willingness of the modellers and users to engage generally limit the number of iterations.

8.11. In relation to the modelling work that was conducted by SPI-M-O the details of the number and types of models that were used, developed and fed into the UK government's response to the pandemic are all available in the SAGE repository and/or supplied by the SPI-M-O secretariat. SPI-M-O did not consider any model or suggestion that was not presented to the committee, either as slides or as a working paper. The majority of more important models have been put on pre-print servers (for example a search for 'SPI-M' on *medrXiv* [GM/190 - INQ000213134] returns 136 accessions (23 April 2023) and have been or are being published (for example, the special issue of the *Philosophical Transactions of the Royal Society B* [GM/191 - INQ000213135]). I am happy to answer specific questions, but am not in a position to do a thorough review, both because of resources available, but also because I was not keeping a formal record of all submissions to SPI-M-O.

8.12. In respect of the CMMID COVID-19 Working Group, the details of the number and types of models that were used are available in the repository. Dr Eggo and Professor Edmunds will be able to answer specific questions.

a. In terms of the number and types of data streams that were utilised, how the data was integrated into the models and how the data streams evolved during the course of the pandemic, this is again a very large task that I am unable to undertake. The data streams were reviewed during the epidemic and it was the SPI-M-O secretariat who held and collated this information [GM/192 - INQ000213136] [GM/193 - INQ000213138].

b. As discussed in paragraphs 8.1 to 8.10, the choice of which variables and assumptions are included in models does not have a formal basis for judgment, so is subjective and dependent on the expertise of the modeller and modelling team. The two approaches that I am aware of to increase the robustness of the modelling evidence are to ensure that the modellers and modelling teams are very expert, and to use the model results in an ensemble. The models used by SPI-M-O were different, and deliberately so. The different modelling groups used different approaches, different data streams and different assumptions to build their models. SPI-M-O secretariat and I made no attempt to guide or influence the modellers' decisions, other than to highlight what the important policy questions were likely to be which, to some extent, determines the minimum structure of a model.

c. I am not aware of any validated approach to measure the willingness of the public to adhere to lockdown over time, or to predict the impact of any particular policy. Different modellers used different data streams to inform behaviour, although the main source of data relating to behaviour was the estimation of transmission rates from case and hospitalisation data, which measures the 'effective contact rate'. We included the potential effect of LD or other policies in the MTPS and scenario models by assuming a change in *R*. Forecasting the epidemic depends on being able to forecast what people will do and what they will do in response to changes in guidance and regulations, which is not possible. There were two occasions in the epidemic when behaviour changed more than was estimated at the start of a LD. The first was 19 July 2021, when restrictions were relaxed and the 'pingdemic' started. The second was 14 December 2021 following the CMO guidance for people to prioritise during the impending omicron wave and leading up to Christmas. Neither of these were foreseen. To the best of my knowledge, no SPI-M-O model explicitly included a measure of the willingness to adhere to LD. Even if they had, the predictions of future behaviour would have been very uncertain, and the associated uncertainty would have to be passed to the decision-maker, so I am not convinced that this would have been useful. Members

of SPI-M-O interacted with members of SPI-B throughout the epidemic, with the underlying aim of being able to incorporate more behavioural insight into the models. Whilst the discussions were informative and transformative in terms of understanding, there needs to be considerably more work done to bring the behavioural psychology and infectious disease transmission dynamics areas together. This is a major undertaking in multi-disciplinary research not possible in the midst of an epidemic, just for time-constraints. SPI-M-O did have access to multiple data streams that related directly to behaviour, for example, CoMix, 999 calls and 111 calls, prescriptions, transport usage, school and staff absences. I had expected that these would be combined using machine learning into a single measure of behaviour, or at least early warning of transmission. However, such an approach did not emerge to the best of my knowledge.

d. All the models used by SPI-M-O, and, indeed, any model, have many uncertainties. Some of these, for example, transmission rate, can be estimated directly from data or through the fitting of the models to data. Others, such as the effect of prior exposure and vaccination (i.e., immunity) cannot be estimated and so require assumptions. For many processes (such as immunity and behavioural responses to the epidemic) the current time was at the forefront of the data – there was no way of knowing what the duration of vaccine-derived immunity was until it was observed for the first time during the epidemic. Parameters such as the infection fatality rate (IFR) were estimated from different sources and the uncertainty around the estimates was quantified. I believe that these uncertainties and limitations were communicated well to the policymakers through the SPI-M-O consensus statements which always contained the principal 'failings' of the models as we understood them at the time. There is a competition between clearly expressing the limitations and uncertainties of the modelling evidence (which derives principally from the models themselves) and the value of the evidence. Striking the right balance evolved during the epidemic. Initially, there was a need within the secretariat to ensure that government understood that the models were not predictions and that the evidence provided by them was not to be 'believed' at face value. As discussed in paragraphs 10.1 to 10.9, model outputs (graphs) can be very persuasive. After January 2021, the limitations of modelling appeared to be better understood. The progression in dialogue between policymakers and SPI-M-O is discussed in more detail in paragraphs 8.43 to 8.67. I am not sure that the general public understood the benefits and limitations of the modelling evidence. Many people appeared to take the view that if the models were not quantitative predictions, then

they were useless. As is frequently the case, the evidence was 'cherry-picked' to support views already held. Those who wished to see government taking more measures to reduce transmission emphasised the scenarios with the highest peaks as they were published, and those who wished the government to intervene less compared the highest peaks previously produced with the data to demonstrate that models were over-predicting and should not be used as evidence. All the evidence was public, although I do not remember seeing a SPI-M-O consensus statement referred to in the media.

e. The sensitivity analyses enable modellers to assess the impact of uncertainties within a model (in particular, parameter uncertainty) on the outcomes of interest. Sensitivity analyses are very frequently performed – I suspect that there are very few SPI-M-O papers that do not provide results that address the principal sources of uncertainty. Whilst quantitative sensitivity and uncertainty analyses are possible, these have greatest value when models are being used in a predictive context. In the majority of cases the sensitivity was reported in the SPI-M-O consensus statements in terms of which factors and processes were particularly important. SPI-M-O also tried to give an indication on how and when the uncertainty could be resolved, for example, additional data sources or time horizons.

f. The models generally produced different quantitative outcomes, which demonstrates that they are different and adds to the robustness of conclusions from the model comparisons. It would have been extremely surprising if models gave the same quantitative results. Of particular interest to SPI-M-O were the qualitative comparisons.

g. The main point of having SPI-M-O as a modelling sub-group with multiple expert members is that the multiple models are compared and discussed, which is in effect a peer-review.

    I. The normal format of SPI-M-O meetings was that the modeller would present their model and results, and be questioned and challenged by the members. When all the models addressing an issue had been presented and discussed, there was further discussion about the comparison and what could be learned from the models. The SPI-M-O secretariat kept notes and were able to ask questions and check that their interpretation of the evidence was correct. Staff from across government were listening into the meetings and were able to raise questions during and after the meeting. I presumed that most organisations

which relied heavily on modelling would have some internal capacity to check SPI-M-O outputs and conclusions against their own. Certainly, NHSE had such capacity and I and other members regularly met their data and modelling teams. Given the very short turn-around times and volume of work, I do not think that SPI-M-O could have done any more to ensure that the modelling produced was as accurate as possible.

II. The Oxford-led RAMP Rapid Review Group (RRG) started 27 April 2020 [GM/194 – INQ000213139]. This was a group of academics that organised themselves to provide this service to SPI-M-O members and proved very useful. Prior to that, if there was a piece of work that was likely to be influential, then my co-chair and others used personal connections to identify suitable experts to ask, for example the review of the paper Estimates of nosocomial and community transmission of COVID-19 in the England completed 14 April 2020 [GM/195 – INQ000213331] [GM/83 – INQ000213312]. This process was very useful, and I am very grateful to the organisers and reviewers who provided this service. The Inquiry will need to ask Professor Julia Gog (Cambridge) if it wishes to understand more about this process. SPI-M-O members were aware of the need to ensure that their work was put into the public domain as quickly as possible. The pre-print service was the most used portal, although others might also have been used [GM/190 – INQ000213134]. Such pre-print servers are well used and understood by the academic community. Subsequently, the majority of important models used by SPI-M-O have been published in the peer-reviewed literature.

8.13. The granularity of models improved as the epidemic progressed, largely driven by the increased granularity of data. In particular, the inclusion of age and place in the infection data meant that these two dimensions were included in all the models that SPI-M-O relied on. These two dimensions are critical for both the ability of the models to capture the important features of the epidemic, and to inform policy.

8.14. The Inquiry asks whether models were able to "adjust the risk of infection or mortality rates to account for the nature of social contact (such as by distinguishing between indoor and outdoor contact)". This question confuses detail with granularity, I think. The question of whether transmission risks or mortality rates vary between indoor or outdoor contact cannot be answered by a model but must be informed by data. If there is no way of measuring the impact of contact location on risks and outcomes, then they

cannot be included in a model. This was a common error early in the epidemic when SPI-M-O were asked to model issues for which the answer will come from data (see paper from SAGE 47 16 July 2020 [GM/196 - INQ000213140]). I suspect it was frustrating for government that SPI-M-O were not able to answer questions to the level of detail that policy questions ideally wanted even though, as modellers, we could give insight into the important aspects (see paper from SAGE 26 16 April 2020 [GM/197 - INQ000213143]).

8.15. There is good theoretical reasoning to think that contact outdoors carries less risk than contact indoors – although the details of ventilation, definition of indoors/outdoors, duration and nature of contact will carry some weight. There was a lack of direct evidence early in the epidemic, and I am not aware of any evidence since the epidemic (see paper from SAGE 24 9 April 2020 [GM/198 - INQ000213144]). Given that data on the rate and place of contact was not available, we were measuring the combined contact-transmission rate from observed incidence of infection. There is unidentifiability in this measurement, so that, for example, 10 people becoming infected on one day could be due to 100 relevant contacts with 10% chance of transmission at each contact or 1000 contacts with 1% chance. Given the lack of evidence, the modellers have to assume an average contact-transmission rate applies to all relevant contacts.

8.16. Even if there was good evidence that outdoor contact was of less risk than indoor contact, the question remains of what benefit would derive from including this detail in transmission models. There are, I think, two good reasons for including. First, if the propensity for indoor/outdoor contact was highly clustered within individuals and highly correlated with important dimensions such as age and risk of disease. If, for example, older people were far more likely to meet outdoors (and outdoor contact is significantly less risky) then it would have significantly influenced the epidemic pattern. However, I am not aware of any data to suggest that this was the case.

8.17. Second, if government had said that indoor/outdoor settings of contact would play a significant role in policy decisions, then it would have provided an impetus for modellers to ask for the data and potentially include this detail in the models. Where we were asked about this level of detail, for example relating to 'bubbles', it was considered but without any data to inform (see SAGE 34 7 May 2020 SPI-M-O paper [GM/199 - INQ000213145]). Had government thought that there was sufficient benefit to, for example, allowing all and any outdoor contact, then SPI-M-O is likely to have

suggested that the policy be implemented in a manner that allowed the impact of the policy to be evaluated and provided data for inclusion in the models.

8.18. In terms of granularity, one aspect of models that was lacking for all models (not just models in the pandemic) is stratification in terms of deprivation. Socio-economic status (SES) is an important driver of all infectious disease transmission dynamics due, *inter alia*, to access to healthcare, ability to make choices about risks, and prevalence of co-morbidities. The importance of SES to the impact of epidemics is well understood, for example [GM/200 - INQ000213146] It is also highly likely that SES interacts with age and other drivers of transmission so that it is itself a dimension along which mixing rates vary. There was some progress in starting to define these patterns during the epidemic (for example SAGE 87 (22 April 2021) [GM/201 - INQ000213147] However, despite this known and observed importance, models have not developed to be able to include SES or other deprivation indices to the best of my knowledge. This is a gap that the modelling community should address.

8.19. I cannot recall being asked a policy-related question concerning SES. For example, we were not asked whether "policy X creates more inequality than policy Y". I cannot recall SPI-M-O being asked any question which we could not answer because of insufficient granularity in the models.

Explainer: Granularity

8.20. The simplest models assume that all people are identical and have the same average experience. So, for example, the risk of hospitalisation given infection is the same for all, and the chance of meeting somebody infected is equal for all. Although we know that people are all different, the differences only matter if they influence the transmission dynamics of the virus, and in particular the risk of infection, likelihood of transmission, or risk of disease following infection.

8.21. The granularity of a model is not the same as the level of detail included in a model or the structure of the model. It is possible to build individual-based models (IBM) that include all individuals as separate entities, but which nonetheless assume they are all identical. Clearly, a model that includes differences between individuals has to have the relevant structure, so that increased granularity implies increased structural complexity, but not *vice versa*.

8.22. Increasing the granularity means complicating the model so that relevant differences between people are included. Most commonly this is done by including groups within models who are assumed identical within the group. So a model that includes age and place, for example, might have people who are 60-65 years old living in Edinburgh as a group, but will assume all members of that group are identical, and have the average risks and contact behaviour specific for that group, which will be different from people who are 15-20 years old in Glasgow. The more such groups, and therefore the smaller such groups, the finer the granularity of the model. The ultimate granularity is to have each individual included in the model with parameters specific to that individual.

8.23. Models can also include details of behaviour, for example, where people mix differentially: schools, places of worship, care homes and personal care. Such detail creates potential transmission links between people and groups of people. Including such detail would increase the granularity of the model if it was used to define the group structure (e.g., 60–65-year-old living in Edinburgh who had domiciliary care).

8.24. There are two good reasons for increasing granularity and detail in models. First, the ability of the model to capture reality – the 'correctness' of the model, and second, the usefulness of the model for policy decisions (see paragraphs 8.1 to 8.10). These two aspects can be the same but need not be. In both cases the data are required at the same granularity as the model to be able to build the models. Data are almost never available to very fine granularity because of identifiability concerns – for example, the ONS guidance on disclosure [GM/202 - INQ000213148] Models have a lower limit of granularity set by the data available.

8.25. The second good reason for including particular dimensions, and making the model granular along that dimension, is because it is highly relevant to policy questions.

8.26. It is also worth noting that increased granularity does not increase the overall ability of models' accuracy or ability to predict. There is a view, perhaps, that including larger numbers of variables and dimensions in models down to individual preferences and behaviours will increase the accuracy of future predictions, but this is not the case. Increased granularity requires more structural assumptions (which increase uncertainty) and more parameters (with associated uncertainty). ***[explainer ends]***

8.27. The role of SPI-M-O is to generate scientific evidence based on transmission dynamic modelling of the epidemic. The key questions for SPI-M-O are what drives the transmission (epidemiological parameters, core groups – see explainer below) and

what determines disease given infection. The models are intended to inform policy so are guided by what policy options are being considered at the time. The age-dependent risk of severe outcomes, given infection, were well established by the end of February (see paragraphs 4.9 to 4.10).

8.28. Focus on impact of school closures was from early February 2020. There was sufficient evidence from China and preliminary modelling that a large epidemic in the UK was possible and the broad strategy was to delay the epidemic. Consequently, slowing transmission was the basis of early discussions [GM/203 - INQ000213149] [GM/204 - INQ000213152] [GM/205 - INQ000213151] [GM/206 - INQ000213150] [GM/207 - INQ000213159] [GM/208 - INQ000213160]. The only previous large-scale NPI that had been given serious attention was school closures [GM/14 - INQ000212124]. No other specific NPI had been raised at this stage other than isolation/quarantine.

8.29. SPI-M-O models mostly included age by the end of February 2020 with schools included as changes in age-related mixing as schools open and close through the academic year. Note that including the effects of schooling children on transmission does not imply that each individual school is represented in a model. I think that the models included age-related impacts adequately from the end of February 2020. Inclusion of school-based mixing patterns (which implies age stratification or granularity) is common in models because school term dates drive seasonal timing of epidemics for the childhood infections that are the target of many vaccination programmes (for example, measles, mumps and rubella). There has been considerable effort focused on gathering the necessary data, although this improved dramatically during the epidemic (see paragraphs 2.9 to 2.12 and section 7).

8.30. The initial modelling effort in early February 2020 was to adapt current models to include a spatial component. I had raised the possibility that there would be differences between across the UK especially in terms of timing that might influence the public reaction to the epidemic, the public health consequences, and policy decisions. Large-scale (e.g., national) epidemics are made up of many smaller, more local, epidemics. The best approach to modelling this is 'meta-population' models in which each locality is modelled separately. The important policy consequence is that the national epidemic would be longer and lower (proportionately) than the epidemics in each locality, so that no individual place would experience the same epidemic and they would be shorter and sharper than the national epidemic. The degree of disparity between local and national epidemics is largely driven by the amount of connectivity between different

localities, for which we had very little information and struggled throughout the epidemic to get 'mobility' data (see paragraphs 7.6 to 7.23).

8.31. The first reference I can find in my emails to the importance of age in determining the profile of disease is 26 February 2020 in which I note that the epidemics of infection and disease will be separate [GM/209 - INQ000213162]. From a theoretical viewpoint, the huge disparity in risk of disease as related to age, led immediately to the suggestion that if the older people could be prevented from becoming infected, then the great majority of the burden of mortality and morbidity of the epidemic would be prevented. This was originally coined as 'cocooning' from the intervention used to reduce the risk of infants exposure to Pertussis (whooping cough) by vaccinating the household [GM/209A – INQ000260627].

8.32. This is also alluded to in the draft NPI document I sent to SPI-M-O secretariat 24 February 2020 [GM/210 - INQ000213163] [GM/211 - INQ000213164], that formed the basis of the NPI table which SPI-M-O discussed 26 February 2020 and is described in the fourth paragraph; this paper was discussed at SAGE 11 27 February 2020 [GM/212 - INQ000213166] [GM/213 - INQ000213167] [GM/214 - INQ000213168] [GM/215 - INQ000213171] [GM/91 - INQ000075777]. I subsequently did some 'toy modelling' and shared with the SPI-M-O secretariat and some SPI-M-O members to try and stimulate further work and exploration of the potential impact of trying to separate infection and disease [GM/216 - INQ000213172] [GM/217 - INQ000213175]. The idea was further discussed at SAGE 13 5 March 2020 (minutes, paragraphs 3, 9 and 18) [GM/106 - INQ000106152].

8.33. There was some discussion on SPI-M-O of how to include long-term care facilities (LTCF) in the scenario modelling. The barriers were that there was very little data available and that it was not clear how it would inform policy. Data on the number and size of residential facilities is publicly available from CQC [GM/218 - INQ000213176] and I first saw the number and distribution sizes of facilities on 2 March 2020. There were about 500,000 beds, which suggested to me that they might contribute 20,000 patients per week to the peak hospital capacity based on a rough calculation of attack rate, hospitalisation rate and epidemic duration; we had been informed by NHSE that the total beds available was 60,000 so this small population would represent a large proportion of beds. However, there were many more data required before LTFC could be included in models explicitly, most importantly the occupancy and duration of residency, and patterns of staffing. In particular, if staff move between facilities,

separate buildings are essential one larger facility epidemiologically. We also suspected that there was significant interaction with hospital facilities (both residents/patients and staff) which would have needed to be included. The CQC data was also England only. The classification of such facilities was also confused in terms of the types and governance – public/private, different resident groups, size and so forth.

8.34. There was also the question of whether LTCF were important in the transmission dynamics in the whole population – they clearly play a key role in translating infections into disease, but it's not clear that they act as amplifiers of infection. It is highly likely that any role they play in, say, increasing the reproduction number is much less than hospitals. Including hospitals in models would be a higher priority, but, again, the data required to parameterise detailed models was not available. The impact of social and health care on transmission during the first wave is addressed in paragraphs 3.97 to 3.109 and paragraphs 6.6 to 6.9. Policy questions to SPI-M-O were at the total population level and at the level of individual facilities and the population level models would not have been greatly improved with the inclusion of residential facilities explicitly. Care facilities were included in the translation of infection to disease, but not in the transmission dynamics of infection.

8.35. Just based on the observations of population sizes, I thought that there was sufficient evidence to inspire considerable investment in protecting such facilities – see for example emails to DFID CSA, DoD CSA [GM/219 - INQ000213177] [GM/220 - INQ000213178]. On the other hand, they are a relatively small group compared to the millions in the general population as shown in the paper referenced in this email so that in terms of proportions it would have made relatively little difference to the total number of people needing hospitalisation which was almost certain to be well over available capacity [GM/221 - INQ000213179]. I remember hearing that the number of people who received domiciliary care both formally and informally is larger than the number of people in residential homes.

8.36. By 13 March 2020 the idea of reducing transmission to vulnerable people had been renamed 'shielding' (largely on the basis of input from SPI-B members on SAGE I recall) – see minutes SAGE 15 13 March 2020 paragraph 18 [GM/222 - INQ000109142]. The theory continued to develop and further extended to the concept that in order to protect the vulnerable it was necessary to reduce transmission to those people who came in direct contact with the vulnerable, for example, carers. This

approach became known as 'super-shielding' or 'extended shielding'. Although a very similar concept, this did emphasise the need to consider the transmission chain that leads to the vulnerable person (see SAGE 50 document [GM/223 - INQ000213181] [GM/139 - INQ000213073]).

8.37. Early April 2020, I was contacted by an economic advisor from HMT, following my interview in *The Times* (see paragraphs 6.1 to 6.5). From my side, I was trying to convince him that, given the huge economic cost of LD, a large investment in shielding was likely to be cost-effective. I personally, and other members of SPI-M-O were concerned that not enough is being done to protect those who are known to be at highest risk if infected (paragraph 16, SPI-M-O consensus 27 April 2020, SAGE 29 [GM/31 - INQ000223519] [GM/32 - INQ000053212]).

8.38. He was clearly thinking about strategies from an economic viewpoint [GM/224 - INQ000213183]. I kept the SPI-M-O and SAGE secretariats informed. The idea of shielding was to reduce the risk of infection to more vulnerable people, given the prevalence of infection in the community, although clearly the greatest reduction in risk is to keep the prevalence of infection low. This approach got developed into 'segmentation', the idea that the prevalence in one group could be allowed to be high whilst keeping it low in vulnerable groups. Whilst the two ideas are related, they are derived from different perspectives. Whilst the idea of being able to segment a population into two separate groups that do not contact each other is theoretically attractive, it is not practically possible, and even if it were, the epidemic of hospitalisation in the less vulnerable group would still be appreciable.

8.39. The issue of LTCF came back to SPI-M-O periodically throughout the epidemic. The modelling of transmission within and between care facilities and interventions to reduce transmission were largely devolved to SPI-M-O member Hall and the PHE epidemiology cell. These were models at a local level, i.e., models of care homes rather than models with care homes in. For example, the email 16 April 2020 in which SPI-M-O is essentially asked to discuss and validate that work [GM/225 - INQ000213184]. This work is discussed in paragraphs 3.97 to 3.109.

8.40. The requirement for including processes explicitly in transmission models is that there is data to parameterise the model, and that there are clear policy questions to be answered. Neither of these were fulfilled in the early stages of the epidemic for LTCF, but both were fulfilled for schools. Nonetheless, the transmission dynamics of COVID-

19 could be characterised and multiple micro-epidemics happening in enclosed spaces (households, offices etc). Those micro-epidemics where the population is more vulnerable (social and healthcare facilities) will contribute disproportionately more hospitalisations and deaths to the total, but it is not clear what roles they might play in amplifying transmission, i.e., whether social care is a 'core group'. This is a question that will need to be answered, but it is a research question rather than a question SPI-M-O could answer during the epidemic. The models would have been improved had they included both social and healthcare. These two settings are intimately linked, and I do not think can be separated in terms of transmission (see for example this paper) [GM/226 - INQ000213185].

Explainer: Core Groups

8.41. All infectious disease transmission occurs disproportionately in 'core groups'. A general observation is that 80% of transmission occurs in 20% of the population. Core groups are different for different diseases. Core groups are composed of people who, because of their behaviour, are at increased risk of infection and have higher transmission rates when infected. The core group does not have to have the highest risk of disease, either individually or collectively, for example, the core group for influenza is school children (they are more likely to be infected and have a higher rate of infection to others) but the majority of severe disease arises in older people who have a higher risk individually and collectively. Interventions that target core groups have a disproportionate effect on transmission.

8.42. Generally, infection is not the same as disease. The policy and public health focus is on disease – had COVID-19 not caused hospitalisation and death, there would have been a very different response. However, the models' focus is transmission as this drives who acquires infection, and therefore disease. ***[explainer ends]***

8.43. My belief is that the potential of modelling to explore the entire duration of the epidemic was underused, before, during and after the epidemic. Note that this is not the same as the interaction of modelling with other disciplines. Paragraphs 3.80 to 3.94 address the issue of the interaction between epidemiology (the direct harms of the virus) and other spheres, especially economics (the indirect harms of the virus and direct harms of the interventions). Regardless of the outcomes being measured, there is a great value in using models to understand the impact of the epidemic accrued over the time of the whole epidemic. In particular, the models can support the setting of objective

goals which then help to define and determine tactical decisions, i.e., those that are local and in real time. The decisions are discussed in paragraphs 11.12 to 11.22, but this question is about overall strategy and goals. In particular, government needs to define what it considers to be a good outcome, or a less-worse outcome given that no outcome is good.

8.44. Strategic direction and goals are important because "decisions will still probably need to be taken at a time when the facts are not fully known." (See Hine report paragraph 3.23 [GM/13 - INQ000213060]), and those decisions can be taken to maximise the chances of achieving goals when there is considerable uncertainty – so-called 'robust' decisions. The range of uncertainty decreases quickly during an epidemic as data accrues, and earlier decisions are more likely to be consistent with overall goals if they prioritise future flexibility. The key to robust decision-making is to know what is trying to be achieved.

8.45. Although military expertise and capacity can be essential in times of crisis, an epidemic is not a war; the virus is not an enemy trying to outwit its adversary. An epidemic is a natural disaster in the same way that widespread flooding or heatwaves are a natural disaster. Managing the nation through such disasters is helped hugely by the fact that it is known that these disasters will occur – there will be another pandemic – and that their dynamics are understood. Having to make dramatic decisions as an 'emergency' measure is a failure to have planned for the inevitable. A particular example is the time required to be able to measure the impact of an intervention. The inevitable delays of disease progression mean that changes in the transmission of the virus cannot be seen immediately, as laid out in the SPI-M-O review of social distancing measures 1 April 2020 (paragraphs 17 and 18) [GM/227 - INQ000213186]. As an example, the introduction of tiers on 14 October 2020 (paragraphs 6.35 to 6.42) had a beneficial effect, and possibly adequate effect in many places, but it was not possible to measure the effect before the LD announced 31 October 2020 and enacted 5 November 2020. Similarly, doubling times from exponential increases are well understood, and so it is possible to be able to estimate the time to a threshold (e.g., exceeding a previous wave) and plan accordingly. An important benefit of modelling should be that policymakers are not surprised by the future.

8.46. From the viewpoint of the role of strategy for managing the epidemic, I suggest that there are four distinct phases: January 2020 – December 2020, January 2021 – November 2021, December 2021 – February 2022, and since February 2022. I will

discuss each of these in turn, although it is the contrast between them that is perhaps most instructive. Note that my evidence is from what I knew, and there may have been discussions of which I was not aware that filled the gaps I observed. I am also not criticising individuals – the overall strategy needs to be agreed by as many as possible, and especially when lives and livelihoods are at risk, needs to have considerable public support.

8.47. The first phase is the subject of most of the questions that the Inquiry has posed to me and is bracketed by the two biggest waves of morbidity and mortality. There was considerable strategic confusion during this period which meant that decisions had to be taken quickly and potentially sub-optimally. Decisions were apparently taken in response to the immediate situation, and the longer-term consequences of those decisions were not given the weight that I felt and feel they should have been.

8.48. The initial strategy was to have the epidemic in a single wave, i.e., get the epidemic over as quickly as possible. For example, the 26 February 2020 paper contains the sentence "The majority of the population would then develop immunity, hopefully preventing any second wave, while reducing pressure on the NHS" [GM/215 - INQ000213171]. The focus of the SPI-M Modelling Summary is the mitigation of the impact of the epidemic on workforce disruption [GM/14 - INQ000212124]. Healthcare capacity, which I think was the major driver of strategy over the epidemic, is mentioned only once in Annex 2. The aim of the summary is to provide government with the modelling evidence that will be available over the first few weeks of an epidemic before the data to enable more detailed and specific analysis becomes available. Once that data is available, more detailed planning and strategy is possible.

8.49. More targeted planning began during March and April 2020. My email to the recently appointed policy co-chair 28 March 2020 highlights the apparent lack of strategy and that SPI-M-O had not been involved in discussing it [GM/228 - INQ000213187]: "The immediate decisions are going to be driven by the longer-term strategy - what is the longer-term strategy, and should we comment on it?". I also sent a similar email 31 March 2020 to CMO and GCSA to stimulate thought [GM/229 - INQ000213188].

8.50. Discussion focused on the path out of LD in early April 2020 and included how the longer-term strategy affected immediate decisions. A small group meeting 10 April 2020 resulted in discussion 11-14 April 2020 and the paper that went into government [GM/230 - INQ000213189]. I did not see the paper again, but it formed the basis of the

SPI-M-O modelling discussed at SAGE 29 (28 April 2020) – paragraphs 14-17 SPI-M-O consensus statement (27 April 2020) contain the same scenarios [GM/31 - INQ000223519] [GM/32 - INQ000053212]. The question of what was "success" occurred throughout the first 9 months of the epidemic (e.g., email to GCSA 7 September 2020) [GM/128 - INQ000213059].

8.51. Choosing between strategies requires thinking through the multiple possibilities for changes in the future options. Key amongst these is the availability of vaccination. I cannot put it any better than "If we knew when a vaccine would be available, this would completely change the perspective on many of these calculations, as one option becomes containing the infection until mass immunisation is possible. In which case delaying strategies become preferred." (SAGE 29 28 April 2020 paper *Assessment of Changes to Lock-Down and Other Controls*) [GM/32 - INQ000053212] [GM/231 - INQ000213190]. A strategy for dealing with the whole epidemic should bifurcate at the point vaccines become available because they fundamentally change the potential outcomes and therefore the optimum pathway. This might have happened, but it was not made explicit to me.

8.52. The SPI-M planning summary contains the Reasonable Worst Case planning scenario (RWCS) that is provided for government to enable the pre-planning. My understanding was that the RWCS was required legally for various branches of government (local and national) to use in planning. I believe that SPI-M-O was involved in creating 6 RWCS including that from the SPI-M modelling summary. I will not go into fine detail about the RWCS as others will likely cover with more authority and insight – my role was to provide advice and support for the modelling teams as the production of RWCS is an internal government function. The strategic focus of SPI-M-O during February and March 2020 was to develop the RWCS for the COVID-19 epidemic. After the first LD, and with the realisation that government did have the ability to control transmission, development of the RWCS became more difficult, because the path of the epidemic depended on the interventions enacted (i.e., policy options chosen), and the policy choices depend on the path of the epidemic. Consequently, the RWCS ended up as a 'negotiation' between Cabinet Office and SPI-M-O, essentially second-guessing what future policy might be. The interaction with policymakers was good, but the outcome was hampered by the lack of a goal. The discussion in the Wales TAG notes of 30 September 2020 express the problem well [GM/232 - INQ000213191] [GM/233 - INQ000213192] [GM/234 - INQ000213193]. My personal belief is that the

management of the epidemic would have been better if there had been some target or goal against which the modelling could work, as summarised in my email 29 September 2020 [GM/235 - INQ000213197].

8.53. The scenarios that models can generate give policymakers an insight into potential futures that can, to a large extent, be chosen between. For example, the 28 October 2020 paper [GM/236 - INQ000213198]. From my viewpoint, documents such as these 'disappeared' into government and did not come back to SPI-M-O members for discussion. There would, I think, have been much to gain from more open discussion. The appointment of Dame Angela McLean FRS as policy co-chair of SPI-M-O was very important in terms of what SPI-M-O did achieve. Dame Angela has an expert understanding of the modelling and was able to involve policymakers from across government. I do not know the details of what she did, but I have a strong belief that it was highly influential in terms of getting scientific (and especially modelling) evidence understood better by policymakers.

8.54. The second period, from my position, was very different in terms of strategy. The discovery of vaccines in November 2020 and the demonstration that vaccines could be used to prevent hospitalisation and death if people were immunised prior to infection meant that there was a clearer objective. This period started with the biggest wave of hospitalisation and death in January 2021 and subsequent LD. Vaccination was rolled out up to July 2021, and then the relaxation of interventions. The four months up to the appearance of the omicron variant were epidemiologically extraordinary, but fell within the strategy set in January 2021.

8.55. The minutes of the SPI-M-O meeting 6 January 2021 record that "…Cabinet Office provided feedback on how SPI-M modelling had been used recently within government to inform decision making and with the committee then discussed important questions that will be valuable to answer over the upcoming weeks and months." [GM/237 - INQ000213199] This was very important for two reasons. First, in terms of SPI-M-O morale, the second wave (with objectively the worst days of the epidemic) felt like a failure, and we had had almost a year of very little direct discussion with policymakers. To have somebody tell us that our work had been important evidence for decision-making and thank us for it was a huge boost for me personally, and I suspect for other members. Second, to give some indications of what decision-makers were looking for as a good outcome gave direction and purpose to the modelling. I think that the SPI-M-O summary of the modelling at the start of the second phase 3 February 2021

[GM/238 - INQ000213200]) demonstrates a different and more confident SPI-M-O because we knew we were doing work that would be useful.

8.56. SPI-M-O were not asked to produce another RWCS, although I presume that one was required and created out of the modelling results. Again, I think that this was highly significant in that it demonstrates that government recognised it was in control of the epidemic and that it used modelling to inform decisions rather than trying to use models to tell them what they would have to deal with.

8.57. Another contrast in this phase was the way that modelling was used as evidence in public communications for decisions. The amount of modelling that SPI-M-O did was huge, but it was not put forward as the principal evidence for decisions, as had been the case in March 2020 and November 2020. The three modelling groups that did most of the 'roadmap' modelling (LSHTM, Imperial College, Warwick) essentially each produced three years' work every month. This is only possible if the data are available and the reason for the work is known. The timing of decisions was known, and had been sufficiently spaced so that the impact of each decision could be evaluated and used to inform the next. The suggestion of adequately spacing interventions was made in the SPI-M-O consensus SD measures 13 April 2020, para. 17, but only done from January 2021 [GM/132 - INQ000213065].

8.58. The policy co-chair of SPI-M-O arranged for a less senior member of SPI-M-O to work within the Cabinet Office. There were two short-term contracts with two people who became 'embedded' into the policymakers: Dr Nicholas Davies (LSHTM, 6 April 2021 – 6 October 2021) and Dr Louise Dyson (Warwick, 12 January 2022 – 31 March 2022). I do not know any details as they interacted within government, but I suspect that they were important in terms of being able to inform government about what models can and cannot do.

8.59. Models are a combination of data and assumptions, and the assumptions had been discussed with policymakers, so the modellers knew the evidence was directly relevant to the decisions. At each decision point, the modelling analysis was made available for scrutiny, but in contrast to previous decisions, was not presented as the main evidence. The interaction between policymakers and modellers during this period was, I think, exemplary in terms of how modelling evidence should be created and used.

8.60. The third phase is the period around the omicron wave. Whilst this period was as frenetic as the first in terms of gathering scientific evidence, the strategy was clearer.

Although never explicitly stated, the interaction with policymakers earlier in January 2021 indicated that government wished to avoid another wave of hospitalisation that required an 'emergency' LD. In autumn 2021 the evidence suggested that there was still considerable uncertainty in terms of the possibility of another such wave, and the modellers were asked to produce evidence to inform the introduction of 'Plan B', i.e., measures to control such a wave.

8.61. The modelling produced as the omicron wave developed up to the end of 2021 unfortunately became the focus of public attention again and was central to the political debates about measures to be taken. I think that this is inevitable when there is very little data and assumptions are more prominent. Note that policymakers never directed modellers in terms of biological or epidemiological assumptions, but did discuss assumptions that related directly to policy choices. I think that this is appropriate and allows the modelling to be independent but useful. Some consideration needs to be given to how to use and present modelling in such circumstances. I do not have any specific suggestions.

8.62. One of the relative failings of the modelling evidence at the start of the omicron wave relative to the previous phase was that the modellers did not know when such decisions would be made. Plan B was announced 8 December 2020 and introduced 10 December 2020 which essentially meant that modelling prior to that date was made obsolete, and did not give sufficient time to measure its effect on transmission. As with the first and second waves, introduction of measures early allows the possibility of evaluating their impact before making further decisions.

8.63. I answered questions on the modelling evidence at the House of Commons Science and Technology select committee 2 March 2022, and subsequently provided written clarification (both documents publicly available on HoC website). I refer the Inquiry to those documents for discussion about the evidence and its role in the policy choices made.

8.64. The fourth phase, since February 2022, is to a large extent the pre-pandemic planning stage for the next epidemic/pandemic and developing the strategy for the next epidemic should be happening now. SPI-M-O has not met since February 2022, although some of the members and I continue to support UKHSA to produce regular MTP. SPI-M has reformed and is meeting regularly to develop evidence to support pre-planning for the next epidemic.

8.65. The SPI-M pre-pandemic summary in 2018 included only one strategy which is to allow the epidemic to happen. Clearly, given the experience of COVID-19, there would appear to be a potential government decision to LD rather than experience the potential for a large wave. The time to discuss this with senior policymakers is now, as it is very difficult early in an epidemic. For example, this conversation (email thread including SPI-M-O secretariat) about potential strategies for dealing with the epidemic should be happening now [GM/239 - INQ000213201]. Determining the multiple different policy choices and trajectories of a potential epidemic is complicated, and will change as societal preferences change and work practices change. So, any policy needs continual revision and updating. The single RWCS does not appear to me to be practical, and the RWCS was dropped from the scientific agenda after January 2021. Although this work is developing it is relatively poorly resourced compared to the impact of a future epidemic, and is very slow paced. Ideally, I would like to see the interaction developed in January 2021 continued rather than seen as a response to an emergency.

8.66. The Inquiry asks about the capacity within government to interact with modellers and use modelling advice to best advantage. I was rather dismayed to be asked to comment on two 'explainer' documents for the prime minister's office and other senior people early in the epidemic [GM/240 - INQ000213202] [GM/241 - INQ000213203]. Whilst I do not expect senior officials and elected representatives to know this, I do expect that there are people within government who are able to explain what modelling is. It was also quite clear with the people that they had not been exposed to modelling previously. I attended a one hour 'teach-in' to Cabinet Office 1 June 2020 with SPI-M-O member Edmunds to introduce people to modelling.

8.67. As indicated earlier in this question, the process changed in January 2021. Prior to that I have already said that modelling was underused. I do not wish to speculate on what changed or why, but it made a very large difference from my perspective. From the perspective of academic co-chair of SPI-M-O, I believe that it would have considerably improved the evidence to support government decisions had SPI-M-O had the same level of interaction from February 2020. Whether it would have made a difference to government's policy choices is speculation.

Section 9: The 'R' Number

9.1. The reproduction number, *R*, measures the number of infections that arise from each infection. If *R* is one, then each infection exactly replaces itself and the epidemic is neither growing nor decreasing. If *R* is greater than one, then the epidemic is increasing exponentially, and the doubling time can be calculated from the value of *R* and the average duration between infections (known as the generation time). The growth rate of the epidemic, *r*, is a very similar measure, although it is easier to calculate directly from data. SPI-M-O also produced official estimates of the growth rate of the epidemic, and translated them to doubling times, largely because doubling times are more intuitively understood. The relationships between different measures of epidemic growth are discussed here [GM/242 - INQ000213210].

9.2. As a measure of the epidemic state *R* has considerable merits and SPI-M-O reported on *R* in other countries (especially China) early in the epidemic, although the key question was whether *R* would be above one in the UK, i.e., whether the UK would experience an epidemic. *R* is determined by the biology of the infection and the patterns of contact in a population and the environment, so it is not correct to think of an "*R* value for COVID-19", and the *R* value for COVID-19 will be intrinsically different in different communities within the UK. However, in the early stages of an epidemic, when infections are either increasing exponentially or decreasing exponentially, the value of *R* is a good indicator whether the epidemic is under control or not. It was also used as a metric of the impact of different interventions on transmission – see for example SAGE 58 [GM/60 - INQ000212102].

9.3. In terms of understanding the current state of an epidemic, and in particular the morbidity and mortality being experienced, *R* on its own is an inadequate measure – the prevalence of infection (the number of people currently infected) is equally important. *R* is a measure of future prevalence so that the combination of *R* and prevalence gives a more complete picture of the state of the epidemic, as explained here for example [GM/243 - INQ000213211]. If the object of control of the epidemic was to keep prevalence low, then that is achieved by keeping *R* low.

9.4. Estimates of *R* can be obtained from multiple data streams – including cases (i.e., diagnosed infections), hospitalisations and deaths. The estimates of *R* are always delayed, i.e., we estimate *R* in the recent past rather than 'today'. The delay is unavoidable and occurs because the actual infection event is unobservable. and the

consequences of infection are observed after a period of time. One big advantage of *R* is that it is averaged, over both people and time, so that it does not 'jump' up and down due to random noise as much as observing the data directly does. In my experience, this is important as observers and policymakers can over-react to short-term changes rather than regarding the epidemic as a longer-term process. The weekly publication of *R* was about right, although two-weekly probably would have sufficed.

9.5. The negative aspect of *R* as an average is that it underplays local outbreaks and requires a relatively large outbreak/epidemic before it can be estimated. When the prevalence of infection is low, the individual chains of transmission are separate but are often insufficient in number to provide sufficient data to accurately estimate an average. At a higher prevalence, the same chains become merged, but there are sufficient cases, hospitalisations, and deaths to enable more accurate estimation of *R*. SPI-M-O members did a lot of work on estimation of *R* at local levels on the spatial epidemiology sub-group.

9.6. With hindsight, using *R* as a policy measure of the degree of control has a significant failing when prevalence is low. Dynamically an epidemic is in the same situation if the number of cases doubles from 10 to 20 as it is from 10,000 to 20,000, but the former is much harder to estimate. This might have played into policy understanding during the summer 2020 when prevalence was starting to increase but there was not sufficient data to reduce uncertainty.

9.7. The question of whether *R* estimate was useful to policy and to public is not one I can answer. It would be worth modellers and policymakers going back over the statistics that were generated to work out what was useful at different stages and for different decisions. For policymakers, I think that the combination of prevalence and growth rate to calculate the doubling time and the number of doublings before reaching a threshold turned out to be the most useful concept. This was the subject of the email sent by SPI-M-O co-chairs 23 October 2020 (see paragraphs 6.35 to 6.42). One of the advantages of exponential processes is that they are relatively predictable, and 'time to event' seemed to be something readily understood.

9.8. I do not know how *R* got to be nominated as the principal statistic for evaluating the epidemic. It appeared in the May 2020 strategy document as an explanation of epidemic process but seemed to grow from there. SPI-M-O produced combined estimates for all nations in the UK and regional estimates in England and several

individual groups published their own estimates [GM/244 - INQ000213212]. SPI-M-O stopped producing an estimate for the UK because it was essentially meaningless, especially if interventions were to be nationally and regionally targeted as described in SPI-M-O consensus 31 March 2021 at SAGE 85 [GM/245 - INQ000213213] [GM/246 - INQ000221772]. UKHSA have since stopped producing official *R* estimates. At a finer spatial scale *R* was a very useful technical concept, but I would not have chosen it as a national indicator.

9.9. As discussed previously, personally, I was not comfortable that SPI-M-O, a group of volunteers, was producing a government statistic which had been given prominence in government strategy, as our role was providing evidence rather than fulfilling operational functions. As a consequence of taking this responsibility on, SPI-M-O had to meet every week to produce the official *R* estimate, without any resource to ensure that people had time away from work, or so that we could manage time for models to be developed and improved. I had a concern that members would either lose interest and stop producing the estimates or that they would all take holidays or decide to recode their models at the same time. The latter concern was mitigated by asking groups to tell the SPI-M-O secretariat if they were not able to produce an estimate for a particular week. In the event neither of these occurred and SPI-M-O members continued to produce regular estimates of *R* to July 2021 when UKHSA assumed the responsibility. A minor operational issue was also created as SAGE stood down from meeting every week, in that SAGE had 'signed off' the R estimates from SPI-M-O [GM/247 - INQ000213215] [GM/248 - INQ000213216].

9.10. I was very keen that Dstl do the final production of the numbers from the multiple models, and also keen that JBC take over the estimates. There was good communication between the SPI-M-O modellers and JBC and then UKHSA – see example emails 3 July 2020 and 5 February 2021 [GM/249 - INQ000213218] [GM/250 - INQ000213219] [GM/251 - INQ000213220] [GM/252 - INQ000213224]. However, it was not until 19 July 2021 that SPI-M-O handed over responsibility.

Section 10: The Influence of Modelling on Policy Decisions during the Covid-19 Pandemic

10.1. As explained previously in Explainer: Infectious disease transmission dynamics and models ((paragraph 3.5) modelling is the best approach to being able to understand

the epidemic process, and is therefore critical in a rational approach to policy-making. Models are tools to aid understanding and not a panacea to resolve policy problems. Generally, models are better able to explain why a policy might fail than they are at defining the policy option which will succeed. Models of infectious disease transmission are not able to predict the future accurately beyond the short-term – they can be accurate if behaviour continues to be average and the same as historical, but when behaviour changes, and especially in response to the epidemic, then the ability to predict the future is lost.

10.2. The SPI-M-O consensus 20 March 2020 para. 10 is "It is not possible to meaningfully model the impact of additional measures at present, due to the huge uncertainty around the current epidemiological picture; impact of interventions already enacted; and impact of the additional measures. We therefore cannot say whether rolling them out nationwide would allow us to avoid ICU capacity being breached in areas outside London." [GM/79 - INQ000213306]. The data at this time was not sufficient to be able to provide reliable models (see paragraphs 7.1 to 7.5). Additionally, there was confusion about the measures that had been introduced and we had no data on adherence that I was aware of. Even if the data had been perfect, there was not sufficient time since the measures were announced on 16 March 2020 to be able to estimate their impact.

10.3. I, and I suspect all members of SPI-M-O, recognise the persuasiveness and potential power of models to influence. Virtually every SPI-M-O consensus statement contains caveats and explanations of why the modelling should not be over-interpreted and taken 'at face value'. The interpretation of the modelling evidence is based on the comparison of results between models, not single graphs.

10.4. The evidence that SPI-M-O produced was passed to SAGE and thence into government, and I had no control on how it was used. There were two immediate situations in which modelling was announced as the primary evidence for entering LD: March 2020 and October 2020. In the first of these, as said previously, when data are sparse, models are largely based on assumptions. Ideally these assumptions are agreed between policymakers and modellers. At the start of the epidemic, data are inevitably sparse (aside from the problems that SPI-M-O had accessing the data that were available). Nonetheless, even without a model, it was clear that COVID-19 was

an epidemic of substantial size, and other countries were going into LD as a result of observing the impact rather than forecasting the impact.

10.5. The second occasion in which modelling was used as the primary evidence for LD was 31 October 2020. As I understand it, some previous modelling that SPI-M-O had discussed ended up in public. During September and October 2020, the epidemic had been steadily growing and there was substantial evidence that the hospitalisations and deaths would reach the same levels as the first wave. SPI-M-O members were not in direct contact with senior policymakers, and we were not being asked to model specific scenarios, but it was clear to me and members that we should be producing evidence for future policy, hence the exploratory graphs. After the appearance of the graphs in the public domain, the decision was taken to use them formally, although I was uneasy and had some involvement in the way that they were presented as discussed in these email threads [GM/253 - INQ000213225] [GM/254 - INQ000213226] [GM/255 - INQ000213229] [GM/256 - INQ000213230]. This episode left some people with the impression that modelling was the major influence on decision-making and policy choices.

10.6. The third occasion when modelling became prominent in the media and political debate about policy choices was around the omicron wave in November and December 2021. At this point the modelling evidence was much better than previously because modellers had both good data and the context for the evidence. Government was more careful about not using the modelling evidence explicitly but directly citing the data (see for example *Press Release Prime Minister confirms move to Plan B in England* 8 December 2021) [GM/257 - INQ000213231]. See paragraphs 11.1 to 11.8 for more discussion of this period.

10.7. I believe that the remit of SPI-M-O was clear, and that of SAGE also. The issue of the scope of SAGE, and whether and to what extent we should have considered economic and social implications, has been addressed directly in paragraphs 3.80 to 3.94. On SPI-M-O we were clearly instructed that economic and social implications were beyond our remit. On SAGE, similarly, we were told that economic implications were not in our remit, although the social implications were discussed to some extent in that they have a more direct impact on health and well-being. As I have indicated elsewhere, neither SAGE nor SPI-M-O had the expertise required to address economic and full social implications of government decisions, and it is the role of decision-makers to make the

judgements regarding different harms. This is made transparent in SAGE 58 minutes [GM/60 - INQ000212102].

10.8. From the SPI-M-O viewpoint, as discussed in paragraphs 8.43 to 8.67, there would have been a huge benefit from more understanding in the first 9 months of the pandemic of what government wished to achieve. From January 2021 onwards, there was a much clearer idea of what government considered a better (or less worse) outcome. I am not sure if the SAGE remit meant that policymakers believed that they could not then discuss the overall strategy with modellers.

10.9. The decision-makers (ultimately Members of Parliament) have a very difficult task when faced with a fast-moving, natural threat. Inevitably, the majority rely on the media for information and understanding, with the inevitable biases and political prejudices that different media have. The media, generally, prefer to focus on narratives about personalities and draw attention to disagreements rather than presenting a consensus of scientific evidence. Inevitably this is taken up by more populist politicians, so that events such as the *Covid-19: Forecasting and Modelling* debate in the Westminster Hall Tuesday 18 January 2022 [GM/258 - INQ000213232] are an exchange of opinions and beliefs about individual's work rather than a discussion of evidence and rational policy-making and decision-making. The major conclusion from the transcript of this session is the poor level of understanding of the formation and role of scientific evidence amongst some MP. Whilst there has been consideration of how to inform the public better, there is perhaps some need to consider how to better inform MP, particular when their decisions are made in emergency situations under uncertainty. I appeared at the House of Commons Science and Technology select committee twice, and would have done more if asked, to support MP understanding of the evidence. There is more discussion of decision-making in paragraphs 11.12 to 11.22.

Section 11: Transparency and Communication of Scientific Advice

11.1. As described in paragraphs 8.43 to 8.67, modelling cannot make policy neutral assumptions. Consequently, the more that modellers understand the policy questions and which decisions the evidence is supporting, the better the evidence is for that decision. In any capacity in which evidence is being provided by technical experts – clinical, legal etc – understanding the 'clients' needs and questions is essential. SPI-

M-O were providing evidence for policymakers, so understanding what policymakers need to know greatly improves the modelling evidence for that purpose.

11.2. During a public exchange on social media in December 2021, I attempted to make the point that evidence provided from SPI-M-O was to inform policy-makers' decisions, not to show all possible outcomes, and whilst there are multiple potential scenarios, only a subset are of interest to the policy-choices that government wishes to make. When making the statement that "we generally model what we are asked to model. There is a dialogue in which policy teams discuss with the modellers what they need to inform their policy" I omitted the word "choices" at the end. Government did not stipulate the methods that SPI-M-O members used to analyse the data or produce the models. Government did not influence the assumptions that modellers have to make when data are lacking, for example on the pathogenicity of the omicron variant. Models must also include some assumptions about behaviour change and how policy will influence it, and the modelling evidence is better if government gives indication of which policy options they are considering so that the models can include them. The modellers are free to include any assumptions they wish. Neither I nor any of the SPI-M-O groups would have produced results which gave results that the government wished to see. SPI-M-O did not "churn out worst case scenarios" but produced impartial evidence to support specific decisions.

11.3. The underlying issue in terms of the evidence is the role of uncertainty. The models and the combination of models produce a wide range of outcomes depending on the scenario. None of the scenarios are a prediction. One of the misunderstandings appears to be that somebody is able to give a probability to any of the scenarios, i.e., which one is more likely. None of the models included a scenario in which the peak of hospitalisation was lower than previous waves. In a sense this is a prediction. In the event, of course, none of the scenarios actually happen. SPI-M-O cannot rate scenarios by "plausibility" as this really would be undermining the decision-makers. Uncertainty is discussed in decision-making context in paragraphs 11.12 to 11.22.

11.4. The comment that somehow a large group of independent academics have the same "negativity bias" is not plausible. I was aware throughout the epidemic that one of the roles of modelling and therefore SPI-M-O was that government were not surprised by any outcome, especially worst cases. The media are more interested in extreme outcomes and tend to give them more prominence.

11.5. My interpretation of what occurred during this period, and my comments becoming prominent was that the scientific evidence got drawn into the political debate about what government should do about the impending wave. There was a political opinion that government should not introduce national restrictions and having seen the modelling being used previously in the epidemic as the justification for LD were seeking to undermine that evidence – as would be expected in a political argument. My mistake was to give an opportunity for the modelling evidence to be seen as 'political' when it is not. The view that somehow government unduly influences the modelling evidence was independently evaluated, confirming that modelling is not used to justify pre-existing government policy, but rather informs the policy choices [GM/259 - INQ000213233].

11.6. In the event, the CMO suggested on 15 December 2021 (This is a publicly available COVID-19 press conference on the 15 December 2021) that people prioritise their contacts and the behaviour change meant that the omicron wave was effectively divided into two [GM/260 - INQ000213234]. The hospitalisations peaked in the week of 5 January 2022 and 30 March 2020 with a relatively high trough week of 23 February. None of the scenarios could have hoped to capture that behaviour change. What the modelling demonstrated prior to the wave was that the wave of infection would be sufficiently large that any reduced pathogenicity would be counteracted. On the basis of the evidence, government decided to introduce Plan B, expedite vaccination and lay plans for temporary emergency beds in hospitals.

11.7. The scenarios produced by SPI-M-O included the assumption of the same generation time of infection as the delta variant as pointed out in the SPI-M-O consensus 15 December 2021 paragraph 29 [GM/261 - INQ000213235]. There was no evidence to support a shorter generation time, but a shorter generation time is associated with a higher growth rate for a given *R* value, so there was some suspicion given the very high growth rates that were being observed. It was given higher prominence in the SPI-M-O consensus 22 December 2021 paragraph 7 & 8 [GM/262 - INQ000213237]. The first academic publication demonstrating a shorter generation time for the omicron variant was published 25 December 2021 [GM/263 - INQ000213238], and the SPI-M-O consensus 12 January 2022 contains the results of further modelling [GM/264 - INQ000213239]. However, this could not have been known during December 2021.

11.8. An Institute of Government comment highlights that the issue is with the wider communication, not with the evidence itself, or necessarily with the way the evidence

is used to inform policy choices [GM/265 - INQ000213240]. The evidence was being criticised by politicians and people with a particular political viewpoint. The situation would have been made easier if senior decision-makers had indicated that they understood the modelling evidence and that they did not influence it beyond making it relevant to the decision.

11.9. I did not have a role in deciding the extent to which SAGE attendance and discussions were made public. Personally, I do not have any concerns about publishing the personnel attending SAGE, and, indeed, welcomed it when it occurred. I do not have any expertise in judging the impact of transparency on public confidence. It is my understanding that when SAGE was sitting previously the attendance was not published at the time.

11.10. The question of Government attitude to transparency and the impact of transparency on public perceptions and confidence is outside of my expertise. My general, wider, experience of working with Government is that confidentiality is the default stance, whereas outside of Government it is increasingly transparency. To some extent I understand that decision-makers need time and mental/intellectual space without external influence and lobbying to make their decisions. In most situations, this does not pose difficulties, but in the epidemic, and especially when uncertainty is highest, the lack of transparency became a subject of much public discussion, almost more than the epidemic itself. In future national emergencies, there will be an increasing need for Governments to make clear what evidence is being considered when critical decisions are to be made.

11.11. In the distinction between evidence and advice, there is no reason for evidence and the generation of evidence to be confidential. I was happy to see the SPI-M-O consensus statements published and encouraged SPI-M-O members to publish their models and results. There are good arguments for advice to be confidential between the advisor and advisee. However, there is a continuum between evidence and advice so that deciding what should be confidential is largely a subjective judgement.

11.12. Hopefully, an outcome of the Inquiry will be how and why various decisions were taken. I am for the most part a member of the public in regard to hearing how and why different policy choices were arrived at, such as the introduction of 'eat out to help out' and not to use precautionary breaks to reduce prevalence or keep prevalence lower for longer.

Key questions are whether there was a clear strategy and whether decisions were consistent with that strategy.

11.13. Note that the outcome of a policy choice does not determine whether the choice is correct or not – it is quite possible to make a choice that turns out well and *vice versa*. The quality of a decision is based on whether it was consistent with the evidence and the strategy. In public health especially, given that different decisions will always change the risks of disease and death for different groups, it is critical that people know why a policy choice has been made. The period January 2021 onwards I think is distinguished by the communication of the desired outcomes and how and why decisions would be taken.

11.14. The issue of the impact of failure to communicate priorities to SPI-M-O has been covered in earlier questions – modelling evidence, especially, is determined by the context and the question being asked. We cannot make policy-neutral models. Priorities were not communicated to SPI-M-O up to January 2021, when the communication improved dramatically, and as a result the quality of evidence. From my perspective the decision-making became more transparent and rational. The impending omicron wave was challenging, but I think that the process was consistent with the strategic framework that was laid out in autumn 2021. The announcement of Plan B before it was needed and enacted was rational given the uncertainty in autumn 2021.

11.15. The better use of modelling ensures that decision-makers are not surprised by outcomes. The precise prediction of the appearance and characteristics of omicron were not possible, but policymakers were warned of the potential SPI-M-O document 14 July 2021 para 16 reads: "If a novel variant emerges and successfully establishes itself, options for intervention may be even more limited. A novel variant would be of concern if it has i) a transmission advantage; ii) escape from immunity (vaccine- or natural infection-induced); or iii) leads to more severe health outcomes, or any combination of these three." [GM/266 - INQ000213241], and the SPI-M-O consensus (3 November 2021, para 3) has the warning: "…in the absence of a new variant of concern…" [GM/267 - INQ000213242]. Policy-making and decision-making needs to be made with understanding that the future is unpredictable and therefore uncertain, and mitigations and alternative strategies are needed. This is particularly true of vaccine, where the strategy and desired outcomes are very different depending on the existence of a proven method of dramatically reducing the morbidity and mortality.

11.16. When communicating uncertainty there is a balance to be struck between saying nothing, and drawing the wrong conclusion from the modelling and misleading policymakers (see explainer below). The SAGE framework for communication of risk and confidence worked well for me. The SAGE risk scale meant that SPI-M-O and SAGE could assign commonly understood probability of events, e.g., that 'event X is highly likely'. It was also possible to assign three levels of confidence to any statement indicating the strength of evidence on which it was based.

11.17. However, I do not know how well it worked for policymakers. Both of these scales require that those producing evidence make a judgement about where to fall in the relevant scales. It might be better to provide evidence at different scales to help communicate the accuracy/precision of the evidence, i.e., 'what is the evidence sure about?' and 'what is the evidence moderately sure about?'. It is not clear to me whether policymakers prefer quite precise conclusions which are more likely to be wrong, or less precise conclusions which are more likely to be correct. My belief is that communication of risk requires discussion so that the degrees of precision and accuracy for different metrics can be jointly understood. It is also my experience (from outside of the epidemic) that policymakers find different metrics more useful than those that come naturally from the modelling. For example, I think that the 'expected time to breaching a threshold of hospital admissions' is more useful for policy than the *R* number (see paragraphs 9.1 to 9.10).

11.18. One potential approach to expressing uncertainty is to produce very large numbers of scenarios that cover many possible futures. The more scenarios that modellers produce the more likely that one of them accurately captures future trajectory. But modellers cannot know which of many scenarios is more likely, and given that the purpose of the evidence is to inform policy the logic behind scenarios becomes circular. It is easier to indicate what will not happen than what will, and SPI-M-O could have made better use of delineating which scenarios we considered extremely unlikely, although this was clear from the MTP and MTPS and accompanying confidence intervals.

11.19. During my evidence to HoC Science and Technology Select Committee in March 2022, Members of Parliament expressed surprise that SPI-M-O and SAGE did not use expert judgement to assign likelihoods to each scenario. Given a set of scenarios that are all possible, decision-makers' focus should then be on the range of outcomes of the scenarios rather than which is more likely to occur. If the modellers or others indicate

that one scenario is more likely, then there is a danger that decisions are taken for that scenario and the other possibilities are ignored (see my written evidence GM/268 - INQ000213243]). Epidemics are particularly challenging, because the epidemic drives the decisions, and the decisions alter the epidemic.

11.20. Such decisions are very difficult and made more so by uncertainty, and it is common for decision-makers to want technical analysts to make the decisions easier by reducing uncertainty. Unfortunately, evidence does not normally make the decisions easier – in fact making the uncertainty clearer will often make the decisions harder, but potentially they will be better. To my understanding, the quality of decision-making is largely about the handling of uncertainty.

11.21. Models are an analytical tool that rely on data. They require work and time, and data flows during the epidemic were continuous. Consequently, modelling results are never completely up to date – some data will always have accrued during the time that the modelling has been done and written up. The prime example was in July 2021 when PHE released updated vaccine efficacy estimates after the modelling to inform the decision about relaxing NPI. This was unfortunate, but inevitable. That modelling evidence is always lagging behind the latest data, and this has to be factored into the uncertainty.

Explainer

11.22. Two useful concepts in understanding model outputs and conclusions drawn from them are accuracy and precision. There is often a trade-off between them, such that conclusions which are precise are less accurate and *vice versa*. For example, the conclusion that there will be a very large wave of infections in the coming weeks might be accurate but is not very precise, and saying that hospital admissions will exceed 1000 per day by the end of the month is more precise but more likely to be inaccurate.

11.23. Uncertainty is a related concept but is (to my understanding) more about the predictability of the system. Some things are inherently uncertain. Prediction of peaks in epidemics is notoriously difficult both in terms of height and timing, and the uncertainty can never be eliminated. Uncertainty can be exacerbated or created by lack of data or knowledge, for example, in March 2020 there was large uncertainty about the state of the epidemic that would have been greatly reduced by better, more timely data.

11.24. The understanding of the relationships between accuracy, precision and uncertainty are well understood scientifically, but not, to my knowledge, well understood in terms of communication to policymakers. ***[explainer ends]***

11.25. The roles of providing evidence, offering advice and decision-making are distinct, and members of the public were confused judging by personal conversations and responses from the public. There was a conflation of the different roles in terms of the regular press conferences at which evidence, advice and decisions were provided simultaneously. The phrase "following the science" has become symbolic, but I am not qualified to judge its full impact. It could be interpreted to mean that decisions will be evidence-based, or it could be interpreted to mean that politicians will do what the scientific advisors suggest. Not surprisingly, I would agree completely with the first interpretation, and be far more circumspect about the second interpretation. My opinion is that if decision-makers follow the evidence from one disciplinary sphere, and ignore, for example, economic impact, then it is a serious error.

11.26. From the viewpoint of the use of modelling, there was a stark contrast between the presentation of the modelling evidence at the announcements of the two first lockdown decisions (March 2020, October 2020) compared to the January 2021 announcement and during the vaccine rollout period (February 2021 to July 2021). The slides and data are available from [GM/269 - INQ000213244]. In the two first lockdown announcements graphs were used from modelling to demonstrate potential futures, but were selections from the full evidence.

11.27. The lack of clarity between evidence/advice and decision-making impacted me most in terms of my interactions with the media, and made my job harder to answer questions about the evidence without it being interpreted as advocating a policy preference or announcing government strategy.

11.28. I am not expert in communication. It is inevitable that messages would change through the epidemic, but this was not signalled well at the outset. There is always a balance to be struck between simplicity and constancy of messaging and updating as required.

11.29. My only strong view on the messaging was the lack of discussion of the impact of the whole epidemic, from start to finish. The announcements were largely about the immediate, or short-term. Although there was some allusion to the possibility of later waves, and the importance of avoiding them, there was no discussion of longer term, for example over three years, and the plans to cope with the epidemic beyond the

current wave. I agree with the view that the messaging swung from over-reassurance to over-alarm as the epidemic cycled, rather than considering the longer-term issues that remained throughout, and would have remained for much longer if vaccines had not been produced.

11.30. In a democracy, and especially in an emergency, the decision-makers are elected politicians – members of parliament and the government they chose. Consequently, the decisions made are political, and epidemiological and scientific approaches will be eclipsed by the prevailing politics. It is a reflection on the nature of politics that a longer-term, more rational approach to messaging and discussion did not occur. The politics will likely be different in future epidemics.

11.31. One aspect that was not explained well in my opinion was the distinction between the individual risks of infection and the population risks of the impact of the epidemic. The focus of government decisions is the population risk of large numbers of people becoming ill and dying at the same time. The interventions and mitigations that government put in place were to reduce the risk of healthcare being unable to cope with the population consequence of the epidemic. The population risk is essentially the individual risk multiplied by the number of people taking the risk. This is alluded to in SAGE 58 minutes, paragraph 4 [GM/60 - INQ000212102].

11.32. In the epidemic, the individual risks are very heterogeneous, but they are taken over a relatively short period and the risks of infection are not independent (as one person's risk of infection is dependent on other people's infectiousness). For most infectious disease risks, governments seek to devolve the assessment and management of the risks to individuals; government seeks to give individuals the information and tools to assess and manage their own risks, but rarely takes the responsibility itself, as, for example, is the case with COVID-19 in May 2023. This distinction between individual and population risks is important in terms of public communication, as most people are concerned with individual (i.e., 'their') risks, whereas SAGE and government are focussing on population risk. The evidence from SPI-M-O was to support government decisions to manage the healthcare load, not to directly help people assess and manage their own risks.

Section 12: Lessons Learned

12.1. From my experience as academic co-chair of SPI-M-O and as an infectious disease modeller there are eight primary lessons that I wish to highlight, in no particular order.

12.2. First, the framework for combining models to create ensembles needs to be carefully considered. There is little question that combinations of specific parameters and outputs benefits from combination. SPI-M-O did that for *R* and MTP/MTPS but rather 'on the fly' and the process that was developed should be evaluated against alternatives. Policy appeared to benefit from derived statistics such as the 'time to a threshold of admissions', and a fruitful activity would be to consider which statistics would be most useful for policy and how to combine such statistics.

12.3. Second, UKHSA (or equivalent) should have full capacity to be able to provide the necessary data and to perform the immediately direct statistical analysis of data streams. To a large extent the analytical capacity was absent during the epidemic, and it was done by SPI-M-O members but also by members of the public who had access to the data and developed novel and important analyses. Making the data public as much as possible was an excellent initiative very well realised in the coronavirus dashboard and other portals.

12.4. Third, I believe that the generation of consensus of modelling evidence from a group of independent modellers worked well. However, if modelling capacity is to be built into UKHSA, then there needs to be thought about how a SPI-M-O equivalent and UKHSA modelling are combined. If there is to be a group of modellers providing independent scientific evidence, then it needs to be adequately provisioned and supported, and not based on academic researchers volunteering their time.

12.5. Fourth, pandemics do not respect national or sub-national boundaries. A global failure was not to have international co-operation and concerted strategies to agree a common approach. When faced with a pandemic individual nations' natural response is to try and gain advantage over other nations which leads to a sub-optimal response from a global perspective. This issue is also, I believe, pertinent to the situation within the United Kingdom where health policy is a devolved issue, so each nation potentially seeks advantage given other nation's policy decisions. The situation in Northern Ireland is particularly complicated and complex given the border with Eire means that a country outside of the United Kingdom has particular influence on the United

Kingdom's epidemic. Having a co-ordinated and concerted approach to the next pandemic would improve strategy development.

12.6. Fifth, pre-pandemic strategy needs to be continually updated. I have pointed out that strategy for how to deal with the whole epidemic, and what is considered a 'good' outcome was not clear until January 2021. The start of an epidemic is not the time to use modelling to go through all the potential threats and the policy response to them; this needs to be addressed prior to the next epidemic. Strategy needs to be refreshed as society changes, so that it accounts for changes in working practices etc. The biggest question is, perhaps, whether LD is to be considered as a future strategy response and under which circumstances, as this has a great influence on preparedness planning.

12.7. Sixth, implied by the previous, is that Cabinet Office and other parts of government that are developing strategy, need to have access to internal expertise to understand how to use modelling to its best advantage. Epidemics should not surprise to policymakers and decision-makers. It is possible, as January 2021 onwards demonstrated, to plan ahead for when decisions will be made and to plan for the data and analysis to be available to support the decision. The presence of a policy co-chair and interaction between senior policymakers (e.g., cabinet office) and modellers are both essential if the SPI-M-O model is to be used again.

12.8. Seventh, government should be better prepared to learn during emergencies. Policies developed through field trails offer a means of more rapidly developing modelling capacity, as well as policy and public understanding. Evaluating policy decisions requires that different places have different policies at different times, which implicitly happened with the introduction of tiers in October 2020, but could have been much more informative if purposively done. If LD is to be an approach to future epidemics, then the epidemics will last longer, and having an adaptive strategic approach will very likely be more efficient and effective. Multiple opportunities were lost during the epidemic to learn, both for the COVID-19 epidemic, but also for the next epidemic. Personally, I think that dissemination of information to help people assess and manage their own risks (i.e., testing and app-based approaches) could be effective and efficient.

12.9. Eighth, this epidemic was remarkable for the amount of data that was generated. The initial experience was not good, but quickly developed. Hopefully, the lessons learned

in terms of being able to gather and disseminate hospital data, and link between different data sets will be built on. Data and analytical systems such as COG-UK and ONS CIS were invaluable to develop understanding as the epidemic progressed and should ideally be part of the preparedness plan.

12.10. The duration of SAGE was unexpected at the outset. The issues relating to the institution and continuation of SAGE are separate. SAGE works across all risks in the national risk register so has to be *ad hoc*, unless there are tens of different standing SAGE committees. The impact of suddenly forming a committee during an emergency could be mitigated to some extent by having permanently standing sub-groups who know each other's work and remit. The GCSA and CMO could then ensure that they had at least met the likely core membership. Whilst SPI-M-O had a working relationship with NERVTAG, for example, there was little interaction with SPI-B, and a working connection with SPI-M-O and SPI-B would have meant that these relationships did not have to form during the epidemic.

12.11. The role of SAGE, in my understanding, is to bring in external expertise and synthesise scientific evidence alongside standing government structures. This is especially important in the early stages of an emergency. SPI-M-O's role in the first few months was largely to assess the size of the problem, although as it became clear how big the epidemic was likely to be, this morphed into working on the mitigations and interventions to reduce its impact. The strategy for coping with the epidemic developed pre-pandemic turned out to be out-dated, and there was a lack of government engagement with SPI-M-O on developing alternative strategies.

12.12. The semi-permanent standing of SAGE and SPI-M-O, to a large extent, filled in roles which Government was unable to fill at short-notice. In 2017 there had been a relatively large modelling capacity in Public Health England, but in the period 2017-2020 much of this was stood down with several senior people leaving and not replaced. I was not aware of any strategic discussion about this, although it meant that by default modelling capacity moved towards relying on multiple models of equal standing rather than a 'champion and challengers' (see paragraphs 8.1 to 8.10). However, the models we relied on were not within government and were largely voluntary (see paragraphs 3.32 to 3.41 for discussion about funding).

12.13. The lack of infectious disease transmission dynamic modelling within government, i.e., done by civil servants, is important. Civil servants have a code of conduct and

conditions of employment which mean that sensitive issues can be addressed, such as the impacts of alternative strategies, without them falling into the political debate. Given that SPI-M-O were a group of external, independent academics I suspect that there was some reluctance to discuss strategy with the wider group as discussed in paragraphs 8.43 to 8.67. The importance of the policy co-chair of SPI-M-O was to develop the trust, understanding and methods of working between those within government and those outside (see paragraphs 3.42 to 3.68). One advantage of SPI-M-O sitting for so long was that the relationship developed to be productive as seen in the evidence from January 2021 onwards.

12.14. The announcement of the setting up of the Joint Biosecurity Centre (JBC) in May 2020 was to some extent a restoration of government's analytical capability, although they largely took on data analysis rather than dynamic modelling. From my viewpoint, the development of JBC and its morphing into UKHSA was slow in that it took over a year before UKHSA took responsibility for production of the official *R* estimate and MTP. My naïve expectation prior to the epidemic was that SPI-M-O would be supporting government capacity and capability rather than fulfilling it.

12.15. I do not have any experience of advisory structures outside of UK, so cannot comment on which countries would provide better models. Key is that whatever external advisory structure is considered is that it must interact with government structures, and it is more important that government structures are adequate given that advisory bodies are small, *ad hoc* and can adapt more easily to the requirements of individual circumstances. My understanding is that the output of SPI-M-O was used by modelling groups in other countries to inform their own evidence.

12.16. The relationship between science and politics is complicated and complex, and although I am aware of a lot of academic work done in this area, I am not expert in it. However, my experience during the epidemic has given me some insights and perspectives that others might not have.

12.17. It seems to me that science and politics are at almost opposite ends of human activity. Politics is based in argument and rhetoric to develop and reinforce beliefs, and science is based on using evidence to ensure that hypotheses are tested. Perhaps the clearest distinction is that "science destroys itself"[1] – new evidence that changes current

[1] Cohen, J., & Medley, G. (2005). Stop Working & Start Thinking (2nd ed.). Taylor & Francis. https://doi.org/10.1201/9780203028148

understanding is heralded as an advance – whereas politics perceives new evidence, which proves former beliefs wrong, as problematic. A major public concern is that governments will present and select evidence to support their argument. The role of scientists is to continually improve current understanding.

12.18. The role of scientists providing evidence and advice to decision-makers is different from normal scientific activity. The role of SPI-M-O was to understand what the policymakers wanted to know and use the best available knowledge to provide relevant information. In normal scientific activity, it is the scientists who decide what the issues and questions are. Providing evidence for policy is more about consolidating current understanding for action rather than developing new understanding. This is more similar to clinical medicine, where patients benefit from current diagnostics and treatment, and the development of novel diagnostics and treatments is to a large extent a separate activity. Nonetheless, the scientific evidence is popularly seen as 'science'. Normal science clearly continued during the epidemic, although many of the advances were technological rather than scientific.

12.19. Decision-making is inherently political if it is done by politicians, which it always is during a national emergency. Politicians are in control of the evidence that they consider, the advice that they take, and the people who create it. Inevitably, during political debate, there is a tendency for the 'science' and the scientists providing the evidence to be pulled into political debate, especially if the decision-makers use the science, or one particular part of the science, as their primary evidence. Personally, I do not think that anybody benefits from poor decisions, and that having the 'science' become part of political debate is more likely to generate poor decisions.

12.20. My strong belief is that keeping the independence and political neutrality of the scientific evidence is critical to maintaining public confidence that decisions are based in reality, i.e., are evidence-based. The SAGE process of generating a consensus of the evidence is fundamentally sound, in my opinion, although it does perhaps need to be more transparent and public-facing, but remembering that it is there to support government primarily, and the public benefit from that support. I would have been content to explain the modelling to more policymakers and decision-makers if requested, although there was not much spare time during the first year of the epidemic.

12.21. The question of how the scientific advice process is subject to a sufficient level of political challenge depends on what the desired outcome is. It is the decision-making that should be subject to political challenge, rather than the evidence. If advice is seen as part of the decision-making process, then the advice should, ideally, be subject to political and public scrutiny. Making a clear distinction between the evidence and advice is a means of keeping the two separated, as it makes it clear where the 'science' ends, and where political judgements are required.

12.22. My experience is that asking the people who are developing the modelling evidence for government during an epidemic to also explain it to the public puts them in a difficult position if there is not political consensus. It is also a significant time burden, and it is a secondary activity.

12.23. Ideally the independent modellers in future epidemics will be available to the media and public, but will not be responsible for explaining the evidence and how it is generated. The explanation ought to be about the policy choices that are being made and the evidence for them. Modelling is part of that evidence, but should not be considered as having supremacy. Other areas of evidence, for example vaccine effectiveness, were explained well, and the same framework might work for modelling.

12.24. The Science Media Centre (SMC) acted as an 'honest broker' for explanation of science and scientific results to the media, and having a non-government organisation involved in communication of science during an emergency of this nature might work well. Note that communication of science and explanation of the evidence is not the same as advocacy for particular policy choices.

12.25. Models (and science more generally) works better when there are specific, answerable questions. Such questions are usually co-created between people of different disciplines, and policymakers also need to be closely involved so that the questions are useful. The framing of such specific, answerable and useful questions was done largely by the secretariats during the epidemics, but it would have been greatly facilitated if the members of different disciplines (i.e., different sub-groups) were already interacting prior to the epidemic. Cross-membership of such groups is, in my view, the most effective way of developing the necessary interactions.

12.26. Interaction with behavioural psychologists (e.g., members of SPI-B) would be hugely beneficial for development of models that could better capture human behaviour. This is best done between epidemics rather than during an epidemic. The usual way in

which such *lacunae* are addressed is through research funding. UKRI and NIHR could do a huge service to the next epidemic by ensuring that the next generation of transmission dynamic models are better linked to behavioural data.

12.27. The Inquiry specifically asks about development of epidemiology-economic models. These exist and are used to evaluate the cost-effectiveness of interventions such as vaccination relatively routinely, and there is relatively good interaction between modellers and health economists. Most importantly, current models do not include dimensions that are not directly relevant to infection transmission, which means that they cannot be used to directly address inequality and inequity. The impact of the epidemic in the UK was exacerbated by disparities in, for example, housing, access to healthcare and occupation. The epidemic and interventions tended to increase these disparities. This is a major gap in modelling and should be addressed. This will need guidance from policymakers in terms of the way in which inequality and inequity are included in decision-making.

12.28. To some extent, the problems faced by government in terms of how to generate, communicate and use evidence to support decision-making during the epidemic were not new. The magnitude and speed of the challenge highlighted deep-seated issues that have long been seen as problematic.

12.29. I am not really in a position to be able to say how it could be improved. Evidence and advice are provided to government for their benefit. Any problems with the evidence and advice provided likely stem from problems with what was being asked for and how it was used, and it is there that the Inquiry should start. I have explained some of the problems I experienced from my perspective, but I was not using the evidence that SPI-M-O generated. As I have said, the modelling-policy interface from January 2021 onwards was a huge improvement and, in some ways, exemplary.

12.30. The scientific research generated within the UK by the funding and personnel was excellent; as examples COG-UK, ONS CIS, REACT, Oxford-Zeneca vaccine, recovery trial were all world-leading in their own way. Exploiting that capability to generate evidence and advice is government's role.

Glossary

CMO: Chief Medical Officer, Sir Chris Whitty

CH: Care homes. This is a blanket term used to describe the whole of residential social care, with particular focus on facilities for older, more vulnerable people.

DstI: Defence Science and Technology Laboratory

GCSA: Government Chief Science Advisor, Sir Patrick Vallance

IDD: Infectious disease dynamics. This is the specialist subject of SPI-M-O and described in paragraphs 3.1 to 3.4.

JBC: Joint Biosecurity Centre

LD: although I do not like the term 'lockdown' because of its lack of established definition and its political symbolism, I do appreciate its brevity. I use this term to describe the three periods of more intense national measures starting March 2020, November 2020 and January 2021. Note that LD is undefined internationally.

LSHTM: London School of Hygiene and Tropical Medicine.

MTP / MTPS: medium-term projections and medium-term projection scenarios – see paragraphs 3.59 to 3.63.

ONS CIS: The ONS coronavirus infection survey [GM/270 - INQ000213245] which provided a huge amount of epidemiological information and gave SPI-M-O a "ground truth" in terms of understanding the epidemic.

NPI: this epidemic has seen the first use of national scale non-pharmaceutical interventions (NPI), and largest scale internationally. Consequently, much of the language and definition is still to be settled. I use NPI rather loosely encompassing all forms of guidance and regulation and different levels of adherence. Clearly, reducing access to schools is a very different type of intervention compared to CMO suggestion that people "prioritise" their activities on 15th December 2021[2]. I have used NPI to describe central, non-personal action (law, guidance etc)

[2] https://www.youtube.com/watch?v=cLS7i1R1fBI at 25mins.

rather than voluntary individual choice, which I again rather loosely call spontaneous behaviour change.

RWCS: Reasonable Worst-Case Scenario.

SMC: Science Media Centre [GM/271 - INQ000213246].

SPI-M-O: the sub-group of SAGE focused on transmission dynamics which I chaired. This is sometimes referred to as SPI-M, which is a separate working group within DHSC which I also chair. SPI-M provides evidence on pandemic preparedness and did not meet whilst SPI-M-O was meeting to provide evidence on the pandemic. Whilst these two committees have similar tasks and cross-membership, they are different.

TTI: Test, Trace and Isolate programme

UKHSA: UK Health Security Agency

Annex B

The CMMID slack channel contains no directly relevant information in terms of the relationship between science and policy. I am not sure who owns it.

The DHSC SPI-M-O slack channels have been archived, and material older than 90 days has been hidden and I have not accessed since the end of SPI-M-O in February 2023. It is owned by DHSC, and I cannot give access.

My emails with relevant, senior people have been disclosed.

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: 4 September 2023