HOUSE OF LORDS

Select Committee on Science and Technology

Corrected oral evidence: The science of Covid-19

Tuesday 2 June 2020

11 am

Watch the meeting

Members present: Lord Patel (The Chair); Baroness Blackwood of North Oxford; Lord Borwick; Lord Browne of Ladyton; Baroness Hilton of Eggardon; Lord Hollick; Lord Kakkar; Lord Mair; Baroness Manningham-Buller; Viscount Ridley; Baroness Rock; Baroness Sheehan; Baroness Walmsley; Lord Winston; Baroness Young of Old Scone.

Evidence Session No. 4 Heard in Public Questions 35 – 43

Witnesses

Dr Paul Birrell, Postdoctoral Researcher, Medical Research Council (MRC) Biostatistics Unit, University of Cambridge; Professor Mark Woolhouse OBE, Professor of Infectious Disease Epidemiology, University of Edinburgh.

USE OF THE TRANSCRIPT

This is a corrected transcript of evidence taken in public and webcast on www.parliamentlive.tv.

Examination of witnesses

Dr Paul Birrell and Professor Mark Woolhouse.

The Chair: I welcome our witnesses, Professor Woolhouse and Dr Birrell, to this session. Professor Hollingworth cannot attend. Thank you for helping us with our inquiry today. Lord Mair will start the questions.

Q35 **Lord Mair:** In the first session we heard a great deal about different models and different assumptions. I think you were both present and able to hear the witnesses' responses. What would you like to say about the different modelling approaches that have been used in the UK in response to the pandemic? Perhaps Professor Woolhouse would like to start.

Professor Mark Woolhouse: Good morning, it is a pleasure to be here. The first thing I should say perhaps is that I was co-opted on to the SAGE subgroup SPI-M at the beginning of this pandemic. My group is not one of the core modellers. We do not produce the outputs of Professor Ferguson, Professor Keeling and the others whom you heard from before. We play more of a support role. I have huge regard for the work that SPI-M does.

In terms of the range of modelling, you heard a very detailed answer on that from my colleagues before, but if we take a few steps back, the picture is a little different. SPI-M stands for the scientific pandemic influenza modelling group, and all these models had their roots in influenza modelling. That is one issue; this virus is not influenza. In fact, it is turning out that its epidemiology probably has more in common in many respects with an earlier coronavirus, the SARS coronavirus. That is one issue, and the models had to be adapted to adjust to this new reality of the Covid-19 virus.

The other issue is what the models have been asked to do. This modelling committee has been set up for a long time and works tremendously well. I have huge admiration for the way it has been set up, but it is set up—or has been used—to address a very narrow set of questions, which largely have been to do with population-wide social distancing: lockdown. It has not been used to explore other possible ways to respond to this epidemic. That is absolutely not the fault of the modellers; they are doing what they are asked to do.

So I think you have uniformity of outputs because the models have similar histories and have been asked to do much the same things.

Dr Paul Birrell: There is a continuum of modelling approaches that are used. As was mentioned in the previous session, they are often based on the susceptible-exposed-infectious-recovered type stratification of the population, but how the models are used often varies greatly. The more complex models, the agent-based modelling that Neil Ferguson would deal with, model every tiny detail of the pandemic.

There are more statistical approaches, which exist to react to data and do not necessarily look at the underlying mechanisms of transmission of the virus but respond to fluctuations in data and will incorporate other data

sources. There is a bit of focus generally on mathematical models and what happens if you tweak a parameter and you make projections about what might happen in the future. There is also modelling—they tend to be simple models—which allow you to incorporate the data and build up an accurate picture of what has happened so far, usually with a coarser stratification of the population. They allow you to come up with robust estimates for quantities that you think are key to the epidemic's future.

Lord Mair: Could you say a bit more about how these statistical models throw new light on some of the modelling that we heard more about in the earlier session?

Dr Paul Birrell: For example, a large agent-based model can be challenged with a relatively small number of parameter sets. You can still consider thousands of scenarios and see what possible future you get. In more statistical approaches you can use a greater array of data, you can model at a finer spatial scale because the models are simpler, and you can get localised parameter estimates where they are available.

There are other considerations there. You are reliant on good local data and on the data not being too noisy, which often happens with local data; they are separate issues. They are of greater utility and are used mostly in short-term forecasting: that is, what will happen over the next couple of weeks. The longer term, more detailed, scenario-type modelling is good for giving good quality statements for what will happen in the future and what might happen with a particular intervention, but for what we can expect in the next two or three weeks, the more statistical models have greater utility, but they do not have the structure to allow you to project into the future.

Lord Mair: Having heard that response, can I ask both of you, in hindsight, whether there are other modelling approaches that could have been useful to aid decision-making on interventions?

Dr Paul Birrell: Interventions tend to be looked at in the longer term. Although policy is constantly changing, no one was implementing a lockdown expecting it to last two weeks, which is the duration for which you might use a model that is particularly good for forecasting. You need more detailed mathematical approaches. What has been done is absolutely right for assessing the potential impact of policies. The mathematical models are essential, and SPI-M focuses on them for assessing the potential impacts of different scenarios, policies and adherence to policies.

Professor Mark Woolhouse: With hindsight, I agree with everything Paul has said about statistical models, but there is an issue about how well you can look into the future using the more detailed models, because you are having to make estimates about what behavioural changes will happen in the future, possibly in response to interventions such as social distancing and so on. You are having to work out how those measures will translate into actual changes in behaviour and what those changes will then do to change the rate of transmission of the virus.

All those things are extremely difficult to do. In the background situation—essentially when people are behaving normally—there are very good data sets, which my colleagues in your earlier session were using, although they are somewhat out of date now, that tell you how people interact and how that might influence the rate of spread of the virus. However, once something like this comes along where, as you have heard, there are not only regulations to change behaviour but a lot of spontaneous behaviour change at individual and institutional level, such as shielding or whatever, we do not know what will happen next, so it is very important when you are trying to look into the future and at interventions that you model a very wide range of scenarios, because that is all they are. They are just possibilities—things that might happen next—not predictions. We cannot predict with this virus.

Lord Mair: Do you think that this wide range of scenarios is already being modelled?

Professor Mark Woolhouse: My group takes a different approach, which does not feed directly into policy—let us call it a reality check. We model hundreds of thousands, and possibly millions, of different parameter combinations to reflect the fact that we are not sure. We do not know exactly what will happen next, so we look at the full range of possibilities and see what comes out that looks robust and what comes out that looks fragile. That is a very important exercise.

Q36 **Lord Kakkar:** To understand in a little more detail the approach being taken to the valuation of models that are having such a big impact on our public health strategies and, as a result of the policy derived from them, on our economy and the broader well-being in society, how does this evaluation take place? What methodological approach is used to determine whether a model that is used in such a profound way is indeed appropriate? How, for instance, is an assessment made of the quality of the data inputs into these models?

Dr Paul Birrell: To answer that question, it is necessary to build on something that Professor Woolhouse has just said. These models come up with potentially many hundreds of thousands of scenarios, but they do not rate the relative likelihood of those scenarios; they are just a list of possible alternatives for what might happen in the future. If we observe something happening in the real world that is different from what the models might have predicted, that is not necessarily to say that the models were performing poorly or were necessarily wrong. All models are wrong, but some are useful, as the famous quotation goes, so it is not necessarily a case of a model being drastically wrong; it might be that the parameter sets against which it is being challenged are not quite right in some way. It is often not so much the model but the parameters that should have primacy in this case.

How are the models assessed? Going back to my previous answer, there are two approaches: the prospective modelling and how we deal with data. Over time, as data becomes available, the models will be used more formally to react to that data and to estimate parameter values. The greatest thing you can do is to contrast the parameter base that you

challenged your model with in the first place when you were prospectively assessing what might happen with what you estimate when you eventually have some data on the outcome that you are interested in.

For example, with much of the early modelling in early February we assumed R values—the more famous parameter that everyone seems to know about—in the range of 2 to 2.4, and we looked at what might happen in those circumstances. However, we now know that R was much closer to 3, that kind of ballpark, which would have had a significant impact if the epidemic had been allowed to run unmitigated. We can see the potential impact and the burden that has been saved by having the lockdown intervention.

Professor Mark Woolhouse: This is a very good question. Verification and validation of the sorts of models that are used in real time have always been extremely challenging. I have worked in this area for 40 years and a recurrent question has been: how do you do that to make sure that your models behave as well as possible?

There are three broad approaches to it. First, before even attempting to use your model looking forward, you ask whether your model captures what has already happened. That is sometimes referred to as hindcasting. You ask whether you can confirm that your model captures that. Secondly, does your model make predictions, and do those predictions come true? Are they consistent with the data? Paul and I have both said that in this particular case that is extremely difficult to do because things are changing so fast, including people's behaviour. Indeed, with the government policy of changing people's behaviour, it is quite hard to do.

The other thing that you can do is look at variables reported in the analyses that you did not explicitly model. To give one example, a very simple model that we ran predicted and modelled the number of people infected in a given time, and it also predicted the number of people exposed. That was not something that we had data on at that time—basically it involved the number of people showing that they had antibodies—but it predicted that.

Then the antibody data came out. I think the first batch was reported by Patrick Vallance at one of the press conferences. He reported a particular figure, so I immediately asked my team, "What's our model saying about last week's level of antibody positives?" They said, "It predicts about 6%", which was exactly what Patrick Vallance said. So we felt reassured that our model had successfully captured something that it had not had any data inputs for. That sort of internal validation happens continuously with all these models.

Lord Kakkar: How often have those approaches been used? The middle approach, we accept, is difficult in a fast-moving situation, but how often have the other two approaches been used—testing the models that we have become so dependent upon for the development of policy?

Dr Paul Birrell: The SAGE subgroup on which Mark and I sit—SPI-M—has a further subgroup on forecasting. Six or seven modelling teams

meet weekly. Matt Keeling and Neil Ferguson lead two of those teams, and Ellen Brooks Pollock works on one of those teams. We provide forecasts of what will happen in the coming few weeks. Independent oversight is provided by a team from Dstl, which measures those forecasts.

We find that different teams forecast different quantities. Different teams are trained to focus on or forecast different quantities well: the modelling that I am involved in seems to forecast deaths well; the model that Matt Keeling is involved in is better at estimating hospital bed occupancy; and so forth. As an ensemble, we can build together a group-think type of forecast. We gather each of the strengths and come up with forecasts that are better equipped based on the totality of our modelling attempts, rather than individual modelling attempts.

Therefore, it is not so much a case of assessing the individual models; it is more a case of combining them and using their gathered strength to see a picture of what will come in the next few weeks and how we appear to be doing at building an accurate picture of the pandemic.

Q37 **Lord Hollick:** The thread running through what we have heard from all the witnesses today is uncertainty. To what extent have the modellers been able to change their approach, and possibly their parameters and the input data, to take account of this uncertainty? The impression left with the public is that we have a scientific, talismanic figure which politicians then use to inform their next policy move. However, we are learning today that there is a great deal of complexity and indeed uncertainty in this. I want to go back to a point that you made, Professor Woolhouse. You said that the use of the data was very much focused on social distancing, which rather begs the question of what it should have been focused on besides social distancing.

Professor Mark Woolhouse: I hope that I said specifically population-wide social distancing—a blanket approach. In the initial stages the models were not asked to focus very much on where the public health burden of this infection lies. It is vastly disproportionate. Three-quarters of all the deaths due to Covid-19 in the UK have been among the over-75s.

Of those 40,000 deaths, three have been in the under-15s. The burden of this disease is vastly disproportionate, so as a worker in public health—as I say, I have been in this field for many years now—in a situation like that I would naturally expect the public health response to be massively targeted where it was needed. It is needed roughly 10,000 times more in the over-75s than in the under-15s. So I would not expect to see a population-wide response of any kind; I would expect to see a targeted one.

So far there has not been so much interest in targeted responses—except in very particular circumstances, which I am sure we will come back to—and the models have not been asked to, and have not really considered, those targeted responses in detail. Shielding, for example, a very appropriate strategy for people who are particularly vulnerable to this

virus, has been part of government policy across the UK nations practically since day one, since the very early stages, but it is not included in any of the models. We are not aiming the models at the right target; we are aiming them at everyone when in fact the burden of this disease is very concentrated.

Dr Paul Birrell: I agree entirely with what Mark says. When looking at the impacts of something like shielding, we were asked once to look at it around February but in the context of looking at population measures to improve shielding, not highly targeted ones. Typically, we react to commissions that come from SAGE, which presumably come in turn from the Cabinet Office, and they tend to be targeted at population-level impacts.

Lord Hollick: In a sense, then, not all the right questions have been asked. Has there been any move among the modelling community to seek to persuade policymakers that there should be a greater focus on some of the sub-population issues that you and Professor Woolhouse have mentioned? Has that been a topic of debate and discussion?

Dr Paul Birrell: One discussion, and there is disagreement here, is on the possible regional application of measures. There is regional disparity. London saw if not a more rapid growth then a greater level of infection prior to lockdown, and correspondingly has had a much more rapid decline since, so there is a question to be asked whether London could have a differential relaxation of measures to elsewhere in the country. The question of age is still not one that I think has been widely discussed. Maybe Mark would like to add to this.

Professor Mark Woolhouse: With regard to age, there is an enormous amount of modelling activity going on around the world. If you go into the pre-print servers, you will find any number of papers on modelling different aspects; some are good models, some are less good models, some you cannot really tell. There has been an enormous amount of interest in trying to model alternative approaches to population-wide social distancing, but it has not yet become what you might call the mainstream modelling direction among my colleagues in the scientific community worldwide. It is more of a niche activity at the moment. We have done lot of work on this in my own group and there is work out there, but as Paul rightly said this is not what has been fed down into SPI-M by SAGE and the Cabinet Office.

Dr Paul Birrell: I have two brief comments on that. Latterly, modelling has been done by a team in Manchester on the epidemic in care homes, but that has not been to assess particular interventions. It has not been tasked with assessing particular proposals or policy changes but is more an attempt to quantify what is going on.

The other age differential that we are asked to look at is the impact of school closures, which is obviously a very hot topic. Again, it is not affecting a very vulnerable group, but it is implementing an age differential policy.

Lord Hollick: Is the lack of testing and tracing data and other data

hindering this more granular, regional, age-based approach?

Dr Paul Birrell: It is. Particularly when looking at granular modelling, you need good data in each sector of the population that you are attempting to model. If you are looking at death data alone, you cannot build up a really reliable picture of what is happening in the under-15s because, as Mark said, there have only been two deaths in that age group, so we do not have the richness of data to infer something useful about what is happening in that group.

Similarly, at a regional level, the data that we have at a fine regional scale tends to be noisy. It might be data on app-based symptom reporting or the number of positive tests in a particular region, such as the PHE line listing of confirmed case data. These data are largely subject to ascertainment effort, particularly the confirmed-case data, so the number of positive tests is relevant to the number of infections in a particular region but perhaps more relevant to the number of tests given. Until recently, when the data on the different pillars of testing were made available, that fine spatial scale did not exist to support the use of the data.

Professor Mark Woolhouse: On the granularity issue, it is very important to understand the limits of the granularity of these models. For example, the question came up recently whether it would it be safe, particularly in terms of R numbers, to open small shops. The models cannot answer that question; they run out of road when you get to that level of detail.

However, the granularity that is there is quite important. There is quite a lot of detail on schools, as Paul rightly says, because they have their ancestry in flu models; school-aged children in schools are very important in the transmission of flu, so the models were very good at coping with them. They did not have care homes in them at all, at first. Where we really needed the granularity for this was not in schools. There is very little going on in schools; there has never been an outbreak in a school worldwide that we know of[1]. Care homes are a different matter, but they were not in the models.

Another very good example of the granularity problem is contacts. Contacts are defined in the models due to people's self-reported behaviour regarding who they see, but it does not say whether they see them indoors or outdoors. There is a massive difference with this virus between indoor and outdoor contexts, which has now been captured by policy; hence the releasing of restrictions on outdoor activities in particular, because it is very obvious that those are far safer than indoor activities. The models do not cover that; they do not distinguish between those two kinds of contact.

Q38 **Baroness Hilton of Eggardon:** To pursue the point about care homes and hospitals, presumably the models predicted the crises in hospitals.

[1] I have since been made aware of two outbreaks associated with schools: one in New Zealand and one in France.

They do not seem to have predicted the crises in care homes, although both Spain and Italy had serious problems with deaths in care homes. I do not know whether you would like to comment on that in view of what you have just said about particularity.

Professor Mark Woolhouse: I am afraid I have not studied in detail the models from Italy and Spain. I can give you some empirical data. I am based in Scotland and am part of the national advisory group on Covid-19 for Scotland. In the first week of the lockdown—I am just reading some figures here—one-quarter of the deaths in Scotland were in the over-85 age group, but now it is more than half. At the beginning of the lockdown, fewer than 10% of them were in care homes; now it is more than 40%. So there has clearly been a differential between what is going on inside and outside care homes. My models do not capture that, and I am not sure that the other models do either.

Baroness Hilton of Eggardon: Does the modelling that is now being undertaken actually assess different approaches to how we might lose some of the current restrictions on people, and what next steps should be used to manage the pandemic in future?

Dr Paul Birrell: As modellers on SPI-M, we are being asked to consider different scenarios. For example, the current programme of school reopening is a modelling exercise that was given a few weeks back. We were asked to assess the impact of the gradual school opening programme over a range of possible responses to the policy. If only a certain fraction of the schools go back, what will that impact be, and if that releases a certain fraction of the workforce to go back to work, what will the likely impact of that policy be? I apologise—I have forgotten the thread of the question.

Baroness Hilton of Eggardon: What steps should be taken, and what models were you using as a template?

Dr Paul Birrell: This process will continue; modelling charges will be fed down regarding future policy steps. All that we ask for as modellers is that, as these things are implemented, we have the time to get some data, even if it is just preliminary data, on what is happening and what the impact of that relaxation has been before the next raft of measures is introduced.

Baroness Hilton of Eggardon: Are you looking at data from other countries as a result of their loosening of restrictions?

Dr Paul Birrell: I think the data from Italy is looked at by some teams, certainly. I have been working with the ECDC on its Europe-wide modelling. We get some feedback. As I mentioned earlier, the parameterising should not necessarily be on the models but the assumptions and the data that are used to underlie those assumptions.

What we put into our models outside what is dictated by the modelling exercise is usually informed by what we can pick up from around the world and what is more relevant. Initially in the pandemic, we were reliant on information on China because that is where the outbreak was, but now we think that that data might not be as relevant to us as what

we are now observing from Italy, Spain and France. Therefore, data from those countries, wherever it seems to be relevant, will inform the assumptions that we make.

Q39 **Lord Browne of Ladyton:** In asking this question, I am conscious of your earlier remarks, Professor Woolhouse, that we should bear in mind what these models that we have been discussing were mainly set up to do. The social distancing intervention was introduced, accompanied by a lot of alleged expert commentary—whether it was does not matter for the purposes of this question. People were surprised that stronger measures had not been brought in, and then relatively shortly thereafter, on about 23 March, in what was described as a dramatic change in direction, we moved quite quickly to a lockdown. The question is whether the data and output from the models backed up the respective timings of these decisions soon after each other.

Professor Mark Woolhouse: That is a very interesting and important question. At that stage in March, we were at the exponential phase of the epidemic in the UK. We have seen that exponential phase in other parts of the world, not least in China, where, as Paul quite rightly said, the best data, and most information at that time, came from studies that were going on at that time in China.

When should you introduce a lockdown? If the measure you are introducing to control the spread of the epidemic is designed to be permanent, unless it is very costly in other ways you should introduce it immediately. That slows the rate of spread. If the intervention you are planning to introduce is a short-stop emergency measure, like a total lockdown, while you work out what to do, the right thing to do is to delay a bit.

So the answer to the question about the right time to introduce the measures depends very much on when you think you are going to remove them. When my group was looking at this—I think some of our work is published in the SAGE papers on this—we were considering an emergency lockdown of six weeks, and we tried to optimise the timing with that in mind. We would have given a different answer—we said that very clearly, because we explored this—if we had been told that it would be 12 weeks, 18 weeks, six months or however long it will be. The answer is different. It is one of those issues where the answer you get depends exactly on the question you ask.

Dr Paul Birrell: Mark's answer is excellent. In terms of theoretical epidemic models, when we are new to the subject we all learn that there is always a right time to intervene. There is the potential that, if you are going to implement measures for a particular length of time, they may lead to a second wave of infection that is more severe than the first.

There was concern initially in this pandemic that, after a release of mitigation measures, you might cause a second wave of infection bang in the middle of the flu season, which would be particularly disastrous because of the burden on the healthcare service. That is what we are always taught: that there is an optimal timing for a particular

intervention. I think that as modellers, we were all thinking, "If we can implement a measure for a particular length of time, when is the optimal time to place that measure?"

I think that around mid-March the tide turned. At a meeting of SPI-M around 13 March, there was a shift from that more theoretical thinking to, "We need to lock down now to preserve the health care services from being overwhelmed". In practice, that has been borne out in that the epidemic had an estimated doubling time ranging from between three to six days. That means that if we delayed by six days, to be conservative we could have had double the demand on the healthcare services. That would be highly localised, but that is perhaps even worse. In the end, the timing one way or the other was certainly not orders of magnitude wrong.

Lord Browne of Ladyton: My supplementary question, with hindsight, is whether the model outputs could have been used to support different interventions or perhaps the timing of interventions, perhaps earlier than when we got to the lockdown.

Professor Mark Woolhouse: Yes, as I have just said: if you ask a different question. If you say that you are going to lock down for six months and you ask when you should you start it, the answer is that you should start earlier. That was very clear.

Dr Paul Birrell: If I remember rightly, Mark presented a paper at a SPI-M meeting in perhaps late February or early March which showed that the gradual introduction of measures would also perhaps be more effective.

Professor Mark Woolhouse: But again, starting earlier. If you are going to introduce them gradually rather than in a block, you start earlier.

Q40 **Lord Borwick:** My question is about the public understanding of these models. Clearly, at the beginning of this communication task in February and March, the task was to persuade the public to take it seriously. Clearly that has worked very well.

Now we have a different problem, which is to get people to understand that it is possible to slowly move out of the very early rigorous rules. People are having some difficulty understanding that. I noticed in the daily press conference yesterday people saying, "Well, it's not moved from level 4 to level 3, so why are we opening up?", as if it were a click and instantly the country changes, when obviously it is much more complicated.

As regards the concentration on R, which appears to me to be an average over London, the Shetlands, individuals in flats, hospitals and care homes—very different groupings all pooled together into the one magic figure of R—are we getting this communication with the public right?

Professor Mark Woolhouse: No, in my opinion.

Lord Borwick: How could we do it better?

Professor Mark Woolhouse: The focus on a single R, as you quite rightly summarised, has been a distraction. I have argued consistently for a long time now that to use a single measure as a metric to drive policy would be wrong. I do not think that it is being used a single measure to drive policy, but the impression is out there that this is a particularly critical number. The original policy goals, which I understood and have therefore been considering in all my work on this, have been to save lives, protect NHS staff and reduce the burden on the NHS as a whole.

R has a loose and imprecise connection with any of those policy goals. One way to illustrate that, as I think came up in the previous session, is that it appears that R over the period of lockdown has come down quite significantly in the wider community but has not come down, and may even for a spell have been above one, in care homes. That immediately tells you that R is not the right thing to look at. It is the wrong way around. If I wanted R to be below one anywhere, it would be in care homes. I think it has been a distraction and has kept our eye off the ball, which is, as I keep coming back to, the population over 70. That is where the public health risks lie. R does not capture that.

Dr Paul Birrell: I do not think the interpretation of R has changed over time, but the usefulness of that interpretation has. When we were in the period of exponential growth, the growth rate was more useful, and the R does not give the rate of growth because it does not tell you how quickly people are being affected. Essentially, at that time the epidemic will grow while R is greater than one and will decrease if that is switched around. But that is really useful when the epidemic is widespread. When the pattern of infection becomes much more heterogeneous in the population, such as the description Mark has just given of the care home epidemic, R loses all utility and we should be looking at different metrics.

Lord Borwick: Would it be possible to invent and publicise an R-type figure for those over 75? Would that be useful as a way of communicating where the problem is?

Professor Mark Woolhouse: I do not have a fully formed answer to that, but my personal view is that we should probably go back to old-fashioned public health and think about it in terms of risk: what is the risk to an individual in this location at that time? Apart from anything else, that is very helpful to the individual concerned, allowing them to make informed choices about how they behave. I am not sure whether the R number helps an individual decide how to behave, but it certainly does not help me.

Q41 **Baroness Blackwood of North Oxford:** I would like to follow up on some of the back and forth that witnesses have had with both Lord Hollick and Lord Browne about the nature of models: that what you get out of them is about the questions you put into them.

Dr Birrell, I think you said that you had been asking questions of your models outside that which was dictated by the modelling exercise. Professor Woolhouse, you have made a number of comments about whether there should have been population-wide questions or targeted

policies on age, care homes or shielded communities. Exactly what is the process of asking you, the modellers, questions of your models, and do you get any say in how those questions should be asked?

Dr Paul Birrell: That is an excellent question. The groups that feed into SPI-M are academic groups. They are free to pursue their research agendas as they see fit, but there are questions and modelling requirements that come down from SAGE, so presumably from the Cabinet Office, that are then fed to SPI-M. It usually models requests that have a very short turnaround time—typically two to three days and often over weekends. Therefore, there is not often the wriggle room to address anything other what is being particularly posed. There is then the opportunity to feed back via SAGE and onwards what we might think of the modelling exercise, but essentially the exercises that come down have a strict deadline to inform a particular policy decision. There is not a whole lot of room for interpretation and for addressing questions other than that which has been put down from on high and from SAGE.

Baroness Blackwood of North Oxford: Professor Woolhouse, you mentioned that your model does not include specific datapoints or parameters that you think should perhaps be included, whether it is age or shielded communities. Are there measures which you think, going forward, should now be modelled to improve our response to the pandemic as we ease the lockdown?

Professor Mark Woolhouse: The short answer is yes, but our understanding of how this epidemic is progressing and how to control it does not need to be based just on models. You can do an awful lot with old-fashioned epidemiology, including things like the contact-tracing exercises. At this stage, in some respects—certainly the fine granularity that we talked about earlier—the models are starting to run out of road and we need basic epidemiology at local levels in different parts of the population, different workplaces and so on. In a way, the focus shifts from the models to the data. We need more and better data.

Baroness Blackwood of North Oxford: If we are shifting from modelling to measurement for a lot of the areas, what role would you say modelling has in that intersection going forward, and how would you improve its role and the way in which models are used for the remainder of this period?

Professor Mark Woolhouse: The simple answer is that we would like the models to be tested and calibrated very closely against the relevant data, relative to what is important to measure at the time. As I say, my feeling over the current period of the lockdown is that there was not enough attention in the models to care homes. We have discussed that many times. Indeed, the data from care homes has been difficult. It has been a source of frustration to my colleagues on the Covid-19 advisory group in Scotland that we have not got all the data we need to make informed decisions. It is not just about models. It is about getting the right data, the right information and, the word you used, measurements.

Dr Paul Birrell: We are finding that we can find areas where the models disagree. It is not always on the quantities that are of particular interest

to a particular question; for example, we are currently discovering that there is discrepancy in the number of estimated infections. The number of cases, symptomatic infections, is easier to get at. The invisible infections are harder to get to. This is what the quantitative data on antibody testing should inform.

We have data from various sources of antibody testing, but usually the data-gathering in those studies is opportunistic. The data have their own biases and weaknesses; maybe they sample only a particular subset of the population. I used data from the National Health Service's blood transfusion service; these people cannot have had symptoms in the two weeks prior to giving blood. You know automatically that you may get a good sample from them, but these people could not possibly have been infected within a certain period of time, simply through their inclusion in the data. There is a demand there which we have identified; we need this kind of data to be able to build a better picture of what is going on.

Q42 **Lord Kakkar:** I have a couple of questions. First, are you content that, from whichever source these questions are coming, the correct questions are being asked of the models and the data you have available, that the limitations are properly understood by those asking the questions, and that you and fellow modellers are in a position to dismiss those questions if the methodology for the modelling you have available to you is inappropriate to address them?

Secondly, Professor Woolhouse, you made a point about internal validity checks. Do you have examples of models where that kind of internal validity testing has been undertaken as data became available and models were discarded as a result?

Professor Mark Woolhouse: To take the second question first, the short answer is no. I have not been over my colleagues' modelling efforts in detail. I do not know what exercises they have taken. What was the first part of your question?

Lord Kakkar: Are you content that there is an understanding about which questions are appropriate to put in the modelling we have available and that the limitations of those models are understood by those putting the questions?

Professor Mark Woolhouse: I defer to Paul on that for the context of SPI-M, because he is more embedded in that, but I will make a quick comment on behalf of the Covid-19 advisory group in Scotland, where I am very satisfied that, right up to ministerial level, there is a very good understanding of what is and is not being said by the models and the other sources of evidence on which they are being asked to design policy. I think that there is a very healthy exchange of information and knowledge in both directions there. Communications in that context are very good. I will leave Paul to talk about SPI-M.

Dr Paul Birrell: If I may first address your second question with a slightly pedantic answer, it is very rare that models should be discarded, as I said earlier. It is more that the calibration or the assumptions used in

the model may need to be revisited rather than throwing the whole model in the bin. Rather than eliminating a modelling group or a particular model, it is more about having to rethink how that model is trained and what information is informing it.

On the first point, in my experience SPI-M will regularly produce consensus statements about what goes forward from our modelling. It summarises the things we agree on and identifies points where we do not agree, which are therefore still uncertain and on which further work is needed. In essence, these statements are usually quite simple. Where there is agreement we can quantify, it is quantified within a reasonable range of uncertainty. I think that line of communication is acceptable.

As we have said before, often the wrong question is asked in the first place, though. We can provide an answer that is understood, but maybe the wrong question is being asked in the first place. We have been asked a number of times to provide a modelling scenario, which meant a single epidemic curve. I do not think that a single epidemic curve is useful, because it will be wrong. You are setting yourself up to fail if you put forward only one epidemic curve, because something will not agree with it in the real world. Particularly earlier on, too often we were prompted by a single epidemic curve going forward for planning purposes. I think it has been a battle to communicate that we should always provide a range of possible values that will hopefully encompass something like what eventually transpires.

Q43 **The Chair:** Thank you very much. Professor Woolhouse, you said that we now need to move from models to basic epidemiology and standard epidemiological metrics, but you are hampered, as I understand it from what you have said, by a lack of data. How do you get this data?

Professor Mark Woolhouse: There is a lot of work going on to design the kind of surveillance and monitoring system that I think we need. It needs to be stratified, just as we agreed—among ourselves, anyway—that perhaps a single R was not the most useful metric in measuring the rate of spread. The same thing is true with the single incidence figure, the single number of new cases per day, and the single figure for the Covid-19 mortality rate. These are all figures that need to be broken up; they need to be granular, by geographic location, age group, risk group, occupation and so on.

We cannot get away from needing to know about pretty much every segment of the population, but standard epidemiology says that you concentrate most of your efforts where you are most concerned. So it is particularly important to get the data in hospitals and care homes. That is starting to happen, but the data management systems we have in place through the NHS and NHS Scotland, which I am more familiar with, are frankly very cumbersome. It has been difficult to extract the right data at the right time for the right person in the right place. There is a lot of difficulty there. That is a historical problem; I have been complaining about it for about 10 years.

The Chair: As I understand it from your answer, you need this to go

forward and manage the pandemic in the future. Is that right?

Professor Mark Woolhouse: The better information we have, with the sort of granularity that Paul and I have talked about, the better we can manage the epidemic in the future.

Viscount Ridley: Could you clear up a contradiction between the first session and the second session? Professor Keeling said in the first session that the models have been "very good at predicting outbreaks", whereas Professor Woolhouse has just told us that we do not know what will happen next and cannot predict with this virus. I do not understand that contradiction.

Professor Mark Woolhouse: I did not hear Matt Keeling's exact words, so I do not know. What can we not predict? We cannot predict the long-term behaviour. Paul said this very clearly at the beginning. We can make fairly simple, statistics-based forecasts for two weeks—I think you said that, Paul, and I would agree. Longer term, there are too many uncertainties. I do not think we can make predictions. Paul gave the right answer a moment ago; in that circumstance, you explore a wide range of plausible scenarios so that you have some idea yourself of what sort of uncertainties you are looking at.

Baroness Sheehan: Following on from your question, Lord Chair, the importance of implementing a targeted response as opposed to a population-wide response has been made quite clear. Do we need to wait for the data before decision-makers can act on that, and how important is it that we act on that immediately?

Dr Paul Birrell: Particularly in the case of care homes, we have the data so can act on that. Generally, without data of any description you cannot really direct your response, because you do not know how to train your model. There is always data available somewhere, there is always something relevant that you can pick up. But, in my opinion, data should have primacy. Data is more important than any model.

Baroness Young of Old Scone: We talked about individual researchers pursuing their research interests. Where should the forum be where researchers get together to talk about modelling strategy? What questions should be asked which No. 10 and the Cabinet might not be asking at the moment? Is there a forum for that where researchers can decide what the questions ought to be?

Professor Mark Woolhouse: The entire academic community, not just the modellers, has been extremely engaged with Covid-19 from the outset. The Royal Society has an independent modelling group looking into questions like that—DELVE, one of its groups, produced a report recently—and earlier this week I attended quite a long workshop at the Isaac Newton Institute in Cambridge.

There are lots of modellers and epidemiologists talking about alternative strategies. The work is being done, but as Paul said it is not necessarily being done in SPI-M.

The Chair: What does all the information about epidemiology and the

model tell you about the likely short and medium-term scenario?

Dr Paul Birrell: We are in a particular state of flux at the moment. With the gradual loosening of lockdown measures, we can no longer predict the behaviour of the public over the coming period, so it is highly uncertain. With each measure that gets relaxed we get closer to a regime of R=1 and around R=1. Very different futures can happen. It is an exciting time, and we will have to wait and see.

The Chair: I do not think it is exciting, is it?

Dr Paul Birrell: It is an excitable system in that what could happen is quite volatile.

Professor Mark Woolhouse: What I think most of the modelling community are concerned about now is that R will increase to something like—let us give a hypothetical number—1.1. I do not think anyone is truly expecting that we will get back to R values of 2, 3 or higher, as they were at the beginning of March. So at this time I would characterise it as our being concerned about a second bump but not a second wave of the kind we saw in the first wave of the virus.

Dr Paul Birrell: The Chinese theory is that an R of 1.1 is difficult to detect, so regimes of epidemics with an R of 0.9, or 0.95—I do not want to be overreliant on R in this conversation—are very difficult to differentiate for a long time, so that is the challenge that is coming.

The Chair: Professor Woolhouse, as somebody who advises the Covid-19 committee in Scotland, I have a great interest in that, as you might well understand. I am north of you. Do you think the divergence that we are seeing is appropriate to the information we have?

Professor Mark Woolhouse: The scientific evidence about this virus is building at an enormous rate; vast volumes of information are coming. It sifted through committees like the Covid-19 advisory group for Scotland, or SAGE for the UK. People are reading it and interpreting it. It is then being passed on to policymakers with those interpretations on them. I am not at all surprised to find that policymakers, not just across the UK but across Europe and the world, are having slightly different interpretations of what they have been told. That will go on.

The Chair: I have overshot my time, but I have not been cut off, which is surprising. I thank both of you for helping us today. It has been a most interesting session, and we appreciate it very much.