HOUSE OF COMMONS

Health and Social Care Committee

Oral evidence: General practice data for planning and research, HC 581

Tuesday 20 July 2021

Ordered by the House of Commons to be published on 20 July 2021.

Watch the meeting

Members present: Jeremy Hunt (Chair); Paul Bristow; Rosie Cooper; Dr James Davies; Dr Luke Evans; Barbara Keeley; Anum Qaisar-Javed; Dean Russell; Laura Trott.

Questions 1 - 86

Witnesses

I: Professor Sir Martin Landray, Professor of Medicine and Epidemiology, University of Oxford, Deputy Director of Oxford University's Big Data Institute, and Joint Lead, RECOVERY clinical trial; Nicola Perrin, Director of Public Affairs, Association of Medical Research Charities; and Phil Booth, Co-ordinator, medConfidential.

II: Lord Bethell, Parliamentary Under-Secretary of State (Minister for Innovation), Department of Health and Social Care; Simon Madden, Director of Policy & Strategy at NHSX and Chair of the Health and Care Information Governance Panel, NHSX; Simon Bolton, Interim Chief Executive, NHS Digital; and Dr Nicola Byrne, National Data Guardian.

Examination of witnesses

Witnesses: Professor Sir Martin Landray, Nicola Perrin and Phil Booth.

Q1 **Chair:** Good morning and welcome to the House of Commons Health and Social Care Select Committee, and our one-off session on the Government's general practice data planning and research programme.

This programme has attracted some controversy in recent weeks with concerns raised about how it has been communicated with patients, the level of anonymisation involved and who exactly will be able to access our data. We know from the pandemic that patient data is vitally important when it comes to things like the vaccine trials and the RECOVERY trial, which discovered dexamethasone, and broader research. The aim of this session is to understand exactly what these changes will involve, how they are going to be used for research and other purposes, what the protections are for NHS patients and what concerns still need addressing.

We will hear from the Government later, in the form of the Minister for Innovation, Lord Bethell; from Simon Madden, who is the director of policy and strategy at NHSX; from Simon Bolton, who is the interim chief executive at NHS Digital; and from Dr Nicola Byrne, who is the National Data Guardian.

First, we hear from both sides of the argument, from those hoping to make use of the data and those who have privacy concerns. A very warm welcome to Professor Sir Martin Landray, who is professor of medicine and epidemiology at the University of Oxford as well as deputy director of Oxford University's Big Data Institute and joint lead of the RECOVERY clinical trial. Welcome also to Nicola Perrin, the director of policy and public affairs at the Association of Medical Research Charities, and to Phil Booth, who is co-ordinator at medConfidential, an organisation that campaigns to protect the privacy of patient data. Thank you all very much for joining us. It is really appreciated.

Can I start with Professor Landray? First of all, thank you very much for the remarkable successes of the RECOVERY trial. When we met last, you said that some people say it saved a million lives over the course of the pandemic. That is a really extraordinary achievement, and we are all really proud of what you have done. Could you outline for us what you see as the benefits of using patient data in the way that some of the changes that are proposed would allow?

Professor Sir Martin Landray: I will do my best. Thank you, Chair, for the introduction. One role that I would also emphasise, which was missed off the list, is that I am a practising clinician. On a Thursday morning, I leave data and trials behind. I close the door and I see one patient at a time in a cardiology clinic in Oxford, so I have direct experience of what it is like to see patients one on one.

In a sense, every interaction with a patient is based on knowledge and experience gained from previous patients, not just the ones I see or that the GPs in this meeting see, but from all patients all around the world over a long period of time. Every time I make a clinical decision—Should this person start a statin to prevent a heart attack? Should they have warfarin to prevent a stroke? How should we treat their diabetes? Might they be at risk of having a fall?—each of those decisions is informed by data and by knowledge from the broader community, if you like, of prior patients and prior experience.

Historically, much of that prior experience and prior knowledge was gained either one by one as a collection of anecdotes, which can be informative but are often merely anecdotes, or by specific consented research studies. Those are important and they have their role. They have transformed much of medicine, but they have their limitations. For example, they have the limitation of size. It is hard to do a very large study on millions of people if you have to go one by one. They have the limitations of variability, the types of people who are included. For example, many of the discussions about who should get a statin or not were historically based on knowledge of 5,000 people who happened to live about 20 miles west of Boston, Massachusetts. We then apply that data to Boston, Lincolnshire, and to every other part of the UK and every other population.

There is a diversity of knowledge that comes from data that can really inform how we predict what might happen to patients; how we understand the causes of diseases that patients might get—whether that is smoking or obesity—and the safety of treatments. What about the safety of vaccines, the safety of common treatments for cardiovascular disease and treatments that might prevent a heart attack? Might they cause cancer? We can understand that from data, and the answer is, by and large, no. The safety of hormone replacement therapy or a breast implant—all of those things—can come from data.

There are variations in care. There is variation in the provision of ophthalmology services. There are variations in diabetes management, in the provision of antenatal and maternity care, and in long-term conditions such as heart failure and chronic obstructive pulmonary disease.

There are variations in time. Are survival rates on dialysis improving or not? Is cancer survival for pancreatic cancer, for example, improving or not? Are our outcomes for bacterial meningitis in children improving or not? In each of these cases data can provide us with the quality of information, the scale of information and the diversity of information to help us make one-by-one clinical decisions.

Q2 **Chair:** That is really helpful. You were able to have the extraordinary successes of the RECOVERY trial without the new general practice data planning and research programme that is being envisaged by the

Government and the NHS. What are the additional benefits of what they are planning to do?

Professor Sir Martin Landray: To be clear, in the RECOVERY programme we linked to 25 different datasets across the four nations. Perhaps many people do not appreciate the complexity. The data are held separately in each of the nations. GP data is held, by and large, in England at least, in two commercially owned and commercially run systems called EMIS and TPP. NHS Digital is a part of the NHS. One of the things that patients complain about clinically, and it makes in some ways little sense, is that we have one NHS badge and one NHS concept of free care at the point of delivery but, bizarrely, we have very siloed data.

To answer your question directly, Chair, with RECOVERY we were able to access 25 different datasets to understand the short-term effects of Covid. We were able to get the GP data under special provision for the pandemic only. We have used that data historically to assess the value of the treatments in different ethnic groups. The best ethnicity data happens to be in the GP record and not in the secondary care record.

Going forward, we will be using the data to study the impact of short-term treatments like dexamethasone, which improves survival, but we have no idea of the impact on long Covid. We have no idea. I think they might actually turn out to be quite impressive, but we need to understand that. Where do people with long Covid go? They go to their GP with breathlessness, with tiredness, with mental health conditions and with cardiovascular symptoms. Perhaps their kidney function deteriorates. All of that happens in primary care. They are the same patients we had in recovery in hospital. They are the ones who survived, thank goodness. We need to understand the impact of dexamethasone on those patients.

Secondly, we need to understand the impact of some of these treatments in people who have long-term immunosuppression conditions. Where does that information come from? That information will largely come from the primary care record.

Both of those will be possible under the existing pandemic arrangements, but I would ask a rhetorical question. Why can we do so much for a pandemic like Covid, which is clearly a disaster with huge health consequences, but we struggle to do exactly the same things with cardiovascular disease, cancer, diabetes and so much else, which are also of pandemic proportions?

Q3 **Chair:** Thank you. I am sure we will come back to you, but I want to bring in Nicola Perrin. You represent the medical research charities. We are going to come to Phil Booth in a moment for some of the counterarguments.

Could you talk through what the opportunities are of using the data from general practice in the way that is envisaged?

Nicola Perrin: Of course. Thank you very much for inviting me here today. Professor Martin Landray has done a fantastic job of setting out the huge range of different uses of patient data that can be made. What was very clear in all the examples he gave was that they were all diseases where our charities are funding research every day to try to improve patient lives. That really depends on patient data, whether it is identifying risk of disease, earlier detection and diagnosis approaches, or to improve treatments.

What the new GP data collection will enable researchers to do is to access GP data on a much larger scale and more efficiently than has been possible before. As you have been hearing, this is not new. Patient data has always been used for research, with safeguards, but this expands the scale on which it is possible to do it. From a patient perspective, it is important that it makes the rules around access much more consistent. It is really improving the governance. It is streamlining things and will, hopefully, reduce the burden on GPs, which is a crucial part of it, so that they can focus their attention on the patient in front of them and not have to worry about the data behind the scenes.

It is really important, as Martin says, that we move beyond just thinking about Covid and remember all the other diseases that one needs data for, whether it is for monitoring the safety and efficacy of treatments or, particularly, for real-world evidence. Using real-world evidence at population scale is going to be the most effective way of really understanding more about multiple morbidities and health inequalities and their impact on different population groups.

We are really excited about this programme. It is hugely important for charity researchers, but obviously it has to be done in the right way.

Q4 **Chair:** Thank you very much. Let me bring in Phil Booth, who represents medConfidential, which campaigns for privacy when it comes to patient data, and to protect the rights of NHS patients with respect to that data.

Phil, can you just talk through some of the concerns you have about this programme? Is there anything you disagree with that you have heard from Professor Landray or Nicola Perrin? What would you like to see changed?

Phil Booth: Thank you for inviting me this morning. As you say, medConfidential advocates for the rights of patients. A large amount of that is around confidentiality, but also on discrimination, which Professor Landray mentioned. We need equity in the delivery of care.

We do not believe that there need be any conflict between good research, good ethics and good medical care. We fully support, agree with and have great admiration for legitimate ethical research done with patients' data. We are not really talking to the benefits that are clearly there; we are talking to, as you say, the concerns that people may have and the concerns that are not, or have not yet been, mentioned, such as other

uses. We have talked about research, but the name of the programme is GP data for planning and research. There is a whole other area of uses.

The way in which data is currently used when it is taken up to NHS Digital is that copies of patients' data are passed to third parties, including commercial entities—some of which have sub-licensing agreements—and there is, essentially, a trade in patients' data. Many of the uses are entirely admirable, but some of them give a large number of people concerns. There are certainly things around commercial reuse of the data, even if it is under some broad, "Well, this is good for health" type—

Q5 **Chair:** When you say "commercial use of the data", all the pharmaceutical companies are commercial set-ups, and yet they come out with life-saving drugs. Are you objecting to that?

Phil Booth: What we object to is the statement, "We do not provide data for marketing," when the service function of the information intermediaries—organisations that get copies of patients' data—is then to charge their customers, which include pharmaceutical companies, data companies and medical advice companies, for market insights and market access. There is clearly a marketing function going on. It is that which gives people discomfort.

Q6 **Chair:** I want to be clear. You are happy with the idea of the data being used for research, but not for marketing purposes. Is that the nub of it?

Phil Booth: We reflect what the public tell us. We understand that there are important functions like pharmacovigilance that are absolutely necessary. Once you have a drug in the market at scale, as we are seeing with the vaccines at the moment, of course you need to track the effects within the population and on individuals. There are also, as you say, companies making money, making drugs and marketing those drugs. What we know from consistent research, studies and polling over years is that many people in the public have concerns about that, whether those are right or wrong. medConfidential advocates for patients and their concerns because it speaks to trust. If we cannot get this right, the big problem will be whether or not people trust what is being done with their data.

Q7 **Chair:** I want you to explain a little bit more because I am not sure that I understand this myself. What use of data are you objecting to by a pharmaceutical company? They are taking data and it is being used in a trial to discover, let's say, a new cure for cancer. You are saying that you do not have a problem with that. There needs to be trust. Patients need to know that their data is being used for those purposes, but providing they know it is being used, you are happy with that. Can you explain the marketing bit that you think patients are not happy with?

Phil Booth: The research and development function of pharmaceutical companies is just one part of what they do. Probably the other witnesses can give a more accurate sense of this, but it is my understanding that

the budget spent on R&D in pharmaceutical companies is a bit smaller than their marketing budget. They have to derive a profit.

Q8 **Chair:** Why would they use patient data for marketing purposes?

Phil Booth: If you are trying to sell your product or service to a buyer—in this country, the primary buyer is the NHS—you need the patient data to demonstrate what your product or drug has done, and that data is being used. We are talking about insights. Again, Professor Landray was very clear about the difference between data and knowledge. Patient data underlies all of this, but the value is derived from the insights.

Q9 **Chair:** You have the data that you are using for the clinical trials. That will then demonstrate whether the new drug works or not. You then share that data with the NHS. At that point, you are not having to share any individual patient data with the NHS, are you? You have the results of the trial. That shows that the drug works or does not work. Is that not right?

Phil Booth: Yes. You do not expose patients' raw data in a published clinical trial.

Q10 **Chair:** I am just trying to understand. Where is the use of patient data for what you are describing as marketing purposes? That is what I do not understand.

Phil Booth: It is because the companies that are given copies of individual-level patient data by NHS Digital, for payment under contract, then service those companies with it. We are down in the detail, down in the weeds. The general public and, clearly, many people do not understand this. It would be better if it was all very clear because at the moment there is concern about the commercial use of data.

We are having a sensible conversation about pharmaceutical companies that most people do not even get to in their conversations. It is not a matter of just—

Q11 **Chair:** I am sorry; I am not sure that I understand the difference. Let me bring in Nicola Perrin, and then I am going to bring in my colleague, Barbara Keeley.

Nicola Perrin: Phil is absolutely right that the public are uncomfortable with the idea of companies accessing health data. The two things that they are often most concerned about are marketing and insurance. The marketing use that they are really concerned about is the idea of targeted marketing to them as a result of their condition, which has been found out using their data. That is not the kind of marketing that Phil is talking about. I think he is thinking more about market sizes and things, which I would argue would be aggregate level anonymous data rather than identifiable data.

By understanding where the concerns are we can make sure that the safeguards address those concerns. The clear red line is that NHS Digital

and the Government have said, "No marketing and no uses for insurance without patients' consent," and really addresses where people are concerned in this space.

Because people are concerned about companies accessing their data, making a profit from it and exploiting data for their own commercial gain, it is very easy to change the media narrative around the use of data to get people worried about a knee-jerk "Any company use of data is bad." From this conversation, it is very clear that it is not bad. There are a lot of extremely beneficial uses where pharma companies need access to data. It is not just pharma companies. That is the other important thing to remember. It is also data analysis companies, software providers, vendors, GP electronic health records themselves, medical devices manufacturers and pharmacies. They are commercial companies too. There are lots of different uses. It is a really nuanced area, but there are a few very key concerns that the clear red lines will help on, as will having strict controls and being able to demonstrate that data is used where there is a public benefit.

There are data access committees that review the proposed use of data to check that the purpose is appropriate. There need to be auditing processes and clear sanctions. Again, they have already been built into the programme. I think that is a really important part of addressing what people are concerned about and making sure that the safeguards tackle that.

Chair: Let me bring in Barbara Keeley. We will come back to you, Phil, because I know that you want to say more.

Q12 **Barbara Keeley:** I want to go back to Phil because I think we need to understand his point about companies being given copies of individual patient data for a charge on the contract. I do not know if you want to say a little bit more about that now, because it will help with some of my later questions. I understand that that would be alarming.

Phil Booth: We are talking about very different things. Nicola has talked about a bunch of protections. Yes, they are notionally there, but saying that you do not sell stuff to insurers or for marketing purposes is not the same as proving it. At the moment, at individual level, linked patients' data is given. Copies of that data are given to information intermediary companies, and they service other companies. We do not know what they are doing with it after that. We do not even have a list of all their customers.

We can talk about a process in theory, but what we can prove is that every month hundreds of copies of vast quantities of individual-level linked patient personal data is—

Q13 **Chair:** Phil, can I clarify? There are people watching this and I want to make sure, as you say, that we do not get too much into the weeds. When that data is anonymised, so it cannot be traced back to an

individual—

Phil Booth: Pseudonymised.

Barbara Keeley: It is not anonymised, Chair.

Chair: Let me finish my point. When it is anonymised—

Barbara Keeley: But it is not anonymised, Jeremy.

Chair: Barbara, can you let me finish my point? I am trying to establish what Phil believes so that I can understand it. If it is anonymised, so that it cannot be traced back to individual patients, are you objecting in those circumstances to that data being passed on to third parties for research purposes?

Phil Booth: We are objecting to what is done. Data is not anonymous. It is pseudonymised. The pseudonym is what allows it to be linked across an individual person's entire medical history. That medical history is a fingerprint unique to an individual, so pseudonymised data is all basically traceable back to an individual. If it was not, it would not be useful for research.

Q14 **Chair:** This is a very important point. Can I bring in Professor Landray and Nicola Perrin to talk about that? Do you agree with what Phil Booth is saying? The data is not properly anonymised, and because it is pseudonymised it can be traced back to an individual.

Professor Sir Martin Landray: I think there have been several things that have been mixed up, as so often in these conversations. Words are slippery. We have already discussed the fact that the word "commercial" is a slippery word and, in a sense, that marketing is a slippery word. There is a huge difference between establishing the size of a market or the types of patients who would benefit most, which of course is what NICE does among other things, and being sent direct adverts for nappies when your family has just had a baby. They are very different sorts of issue. I think we have to be really careful about getting underneath the words to what the concerns actually are.

On the commercial side, particularly the marketing side, I have not really fully grasped what the objections are beyond, clearly, for insurance purposes without consent and for direct-to-consumer advertising. I am struggling otherwise with where the commercial marketing piece is coming from.

On the anonymisation thing, the question is, what does one mean? It is a bit like genomes. We all have a genome. By and large, notwithstanding identical twins, it absolutely identifies us. If you walk along a street and you see a genome lying on the floor and you pick it up, you have absolutely no idea who it belongs to unless you have some other reference information and the willingness to join the two—which is, by the way, in the context of data, illegal under the constraints that are put in place.

Can one in theory, as has been put out by others, identify somebody famous from the datasets, such as a former Prime Minister whose heart operation is widely known and previously reported in *The Guardian*? Yes, one could. Is that legal? No, it is not. If one looked at a very large number of records, could one find people? No.

Q15 **Chair:** Thanks. I want to bring in Barbara because she was doing the questioning. Can I bring in Nicola Perrin and then we will go back to Barbara?

Nicola Perrin: I think this is important because one of the things that confuses people the most is when one party says, "We have anonymised the data. It's fine; don't worry about it," and on the other hand you have some very sensible academics saying, "Nothing is truly anonymous. Don't believe them if they tell you it was anonymous."

The terminology is really important. Understanding Patient Data has done quite a lot of work on how you can really explain this to people. Rather than words, it is easiest if we think about it with pictures. A photo of somebody is clearly identifiable. That is their personally identifiable information. A blurred crowd of silhouettes where you cannot see anyone is fully anonymous. That is aggregate data; it is how many people use asthma inhalers in this country.

It is the grey area in the middle where there is confusion. It is like blurring a photo. The more you blur it, the less easy it is to identify somebody. People recognise that, if you put in enough time and resource and had enough other bits of information, you would be able to identify someone. The really important thing with that blurred image is to make sure that you are not blurring it so much that it is completely useless to a researcher, because the more you blur it, the less useful it is, but you need safeguards around that blurred image. That is pseudonymised information, as Phil has been saying, but it needs protections. That is why those protections are there. With enough will, somebody could re-identify somebody, but it is illegal to do so. You have to reduce the risks as much as possible, while also making sure that you get the benefit and value of the data for the patients.

Q16 **Chair:** Phil, you want to come back. I will then hand over to Barbara.

Phil Booth: Very quickly, let's please not talk in metaphors. The law since 2018 says that pseudonymised data—de-identified data—is personal data and therefore that is the case. Yes, there are sensible things that you can do with data to minimise the risks of identification, but it is not just celebrities. If you have access to the HES data, with just the commonly shared birth dates of their children, you can probably identify a good proportion of women who have given birth to two and certainly three children in an NHS hospital.

There is long-standing research on re-identification. Frankly, saying that it is illegal is no real defence. The audits that NHS Digital performs on

some of its customers show that they have either breached, in some cases, the agreement—the contract—they have with NHS Digital, or in some cases the law. We republish those spreadsheets, which are published every month—the data release register—on a website called theysolditanyway.com. This is NHS Digital data, and NHS Digital audits showing that people are getting this type of data and not following the agreement and, in some cases, not following the law.

Q17 **Barbara Keeley:** These are a couple of comments, really. I am getting a sinking feeling. I was on the Committee in 2013-14, as was Rosie Cooper, when we handled the mess that was care.data. It really appears from some of what Phil is saying that we do not appear to have learned enough from that disaster.

That is my first question to Phil. Why have we ended up in a situation again where privacy is pitted against data use? I understand quite a lot of the issues. In the first part of my career, I worked for the information technology company IBM. I understand data uses and I understand how important they could be. I also understand that when we talk about GP data, it includes sensitive issues like alcohol abuse, domestic abuse, sexual health, HIV status, mental health and things that people do not necessarily want shared. Clearly, we cannot trust certain elements in society not to connect and de-anonymise data, if that is possible, and it sounds like it is possible. What are your first thoughts about that, Phil? Why are we ending up in this situation again?

Phil Booth: Thank you for your question. It is difficult to speak to other people's motivations, but it is clear to all of us that there are major benefits that could be had. The reason why we are in this situation is that lessons were not learned from the care.data process. That went on for two years. There was a big outcry in 2014. The Chair of this Committee instituted an independent care.data advisory group, in which all the various stakeholders agreed that they needed to write to every patient and then spent two years trying to work out what they were going to say and to ensure that what they were going to say was true.

That that process came to an end is by the bye, but the learnings from the process were clearly not adopted when this new programme was begun because the same mistakes have been made again. In fact, in some ways it is worse. Care.data was limited. The amount of data that was going to be taken was clearly defined in a dataset and explicitly excluded many of the sensitive codes and items of information that you have just referred to, whereas this GPDPR extraction is essentially, at first blast, a person's entire GP medical history—the coded information about their treatment, symptoms and all those sorts of things. It is not the free text that the doctor writes or the images that are scanned into the record but, basically, the medical history, and it has not put exclusions on some things. All health data is sensitive, but some areas are considered particularly sensitive and possibly even taboo in broader society.

It is mystifying to me, frankly, why we are here again, when we thought we had made these points clear enough. Nicola Perrin was involved around that period in a different role. Why are we here again?

Q18 **Barbara Keeley:** I want to ask you a few questions first, and then I will come to others. The sinking feeling I have comes from what you said, Phil, in terms of companies being given copies of individual patient data for a charge under contract, and then audits show that they are not doing what they should with that.

The letter to GPs from the Department has only just been issued. It commits that access to GP data will be only via a trusted research environment. Can you explain the difference with that and how that would work? Would that be the solution? It is alarming to think that copies of data are being sold. The whole thing then goes out of control, and that is exactly what we had back in the care.data days.

Phil Booth: Trusted research environments—what is called the Five Safes setting in the Office for National Statistics model—are indeed a way to deal with many of these questions. They stop the need to disseminate or send out copies of the data. We welcome it, and if you look back at the parliamentary record, we were asking for a safe setting to be the approach back in 2014.

It could be a major move forward instead of sending out copies of the data in the form in which it is currently sent out, which introduces a whole bunch of risks. It only takes one person to think that the rules do not apply to them, and even just for a few hours, that would cause absolute chaos. If the data is in the safe setting, in the trusted research environment, and everyone who wants access to it has to apply for that access, be approved and be audited as using it while they are using it, and all the stuff that they take away is checked to make sure that it is not disclosing any individual's details, we could actually not just say that we are putting safeguards and protections in place but prove it.

We very warmly welcome the letter yesterday that said that it would be trusted research environment only. As we were saying, if a pharmaceutical company were to come for a purpose—researching a new drug—that would probably be like anyone else researching anything else. Clearly, that would meet the criteria. If they were coming for a marketing purpose, that could be there as a criteria and you could see what they mean.

The problem at the moment is that different bits of the NHS do different things and have different rules. While I do not believe that NHS Digital knowingly sells data to marketers, we know that pharmacists have done so, and that insurance companies have tried to make patients give what is called an enforced subject access request. We have had to fight all sorts of fights in and around the NHS about individual marketing, which pollutes the public perception.

Q19 **Barbara Keeley:** I understand that. I want to come on to opt-outs because medConfidential has been quite involved in that, and it is a very complex process. It was down to medConfidential to put the opt-out documents on your website page. I think that was a good service.

I believe that every patient should be given full and fair information about patient data. I am sure you agree, but why is that important? Sending letters out to every patient to say, "This is what is being proposed and this is how we are going to use it, what it is being used for and these are the safeguards" seems to be resisted.

Phil Booth: Again, that is something that mystifies me. There have been enough trust-eroding errors made around communications that it seemed obvious to us from the start that you simply wrote to every patient. A letter sent to you, by name, from the NHS, as the NHS, will far more likely be read. For people who do not read it, a copy of that letter can be published. We can just be talking about the letter, rather than a communications campaign that is trying to explain a bunch of stuff, including slippery words, as Professor Landray said. That is very difficult.

In the attempt in May, there were some websites. There was a little bit of tweeting. There was a video—again. There was a fact-checked page that was so misleading that it had to be taken down. We need to make sure, as best we can, demonstrating best effort, that we have put all of these things in place: the trusted research environment; fixed opt-outs; being very clear about all the rules and making sure that is all done and seen to be done. Then we can communicate to patients. Frankly, at that point, we have the opportunity to ask anyone who has opted out if they are satisfied and, if they are now happy with the arrangements, to opt back in. If you keep on doing this other stuff, we are just going to push more opt-outs.

Chair: Thanks. Last question, Barbara.

Q20 **Barbara Keeley:** What is your estimate of the numbers opting out in June, Phil, when there was a flurry of opting out?

Phil Booth: I think that Lord Bethell can speak to that directly. He gave an answer in the Lords to, I think, Lady Cumberlege of a little over 3 million in total. To us, that looks like getting on for a doubling of the number that had previously opted out. At this point, NHS Digital is only able to see what is called the national data opt-out—the online one—because the other type 1 opt-out that you make to your GP is invisible to them. We think that the numbers are roughly the same for both types.

Chair: Thank you.

Q21 **Dean Russell:** Professor Landray, regarding some of those points, how would you address some of the concerns that have been raised? I would also like to ask about where you think the next 10 years are going if we do not do this, in terms of being able to cure cancer and other things.

I refer Members to my entry in the register of interests and to articles

that I have written over the past year or two, being very pro digital health and use of data. I would be keen to get your views as a rebuttal, I suppose, to those points, please.

Professor Sir Martin Landray: Thank you very much. I do not think this is at all like where we were in 2014 with care.data. You can perhaps ask Nicola Perrin that same question. As Phil pointed out, she had a different role at that time.

Why don't I think that? I think NHS Digital as an organisation is substantially more professional in how it deals with data requests and data distribution in all the areas we are talking about. There are detailed application forms that have to be made. It has an independent group advising on the release of data, chaired by a lay person. It has one other lay member, as well as a health ethics expert, an epidemiologist and three GPs.

In the new scheme, there would be a professional advisory group—essentially, a GP group—in addition. There is already an independent group advising on the release of data. They meet every week. In fact, currently it is twice a week, I gather, because of Covid. They publish all their minutes. NHS Digital publishes a log of all the distributions of data that are made. All of that is quite, quite different to where things were in 2012, 2013 and so on. A lot has changed.

Could it be better? Yes, substantially. To find the information you have to be looking for it. It is there all on the website, but it is not the most accessible information. I think there is additional work that could and should be done. Communicating it has been challenging. I think that the communication over the last three months has not been anything like it should be. It will be interesting to hear from those who are closer—from the Government side, if you like—about their reflections on communication, the lessons learned and plans going forward. I want to see that substantially better.

It is worth drawing a comparison. What we are talking about today is GP data. One of the points that was made was that it contains terribly sensitive information. Yes, it does, but we should think very carefully about making that statement as if it is a bad thing.

Depression is a terrible condition. It is potentially a sensitive medical issue. Stillbirth is a terrible thing. It is a potentially sensitive medical condition. Sexually transmitted diseases are certainly sensitive. If all those data were excluded, what would be the impact? The impact would be a loss of ability to find new treatments, manage those conditions or develop the healthcare services that are needed in order to treat those conditions.

I am reminded of the HIV epidemic, as it was considered at the time, back in the late 1980s and early 1990s. Exactly these sorts of issues were considered so sensitive that there should not be data sharing, and that

there should be substantial cautions around HIV tests and so on. Actually, it was the HIV community who turned round and said, "Hang on a minute, you're actually discriminating against us and excluding us from the opportunity for our needs to be met." One needs to balance some of those issues. Just saying that a particular issue is taboo—I think that is the word that was used—is a judgment call that I do not think anybody on this call or in this meeting today is in a position to make for themselves.

Q22 **Dean Russell:** I want to ask Dr Perrin the same question. What is your take on the statements earlier, but also what is the risk of not doing this?

Nicola Perrin: I agree with what Martin said. This is not where we were with care.data because of all the safeguards that Martin has just been through. I think it is wrong to say that lessons have not been learned. The national data opt-out programme, when it was implemented in 2018, demonstrated very thoroughly that the lessons had been learned. There was a lot of involvement with patients, testing, retesting and getting the wording absolutely right, so that people really understood the information that they were being given so that they could make an informed choice.

We were in a very different place. What has gone wrong this time? The obvious thing is not enough communication. That is really frustrating because at the beginning of last year I was in a meeting with people from the programme team and NHS Digital, planning a programme of comms and engagement over a nice length of time, to really take people with them in co-development. That stopped because of Covid. The whole thing was put on hold.

I think Covid has also had an impact on what then happened with the comms. Partly, time had been lost and, because Covid demonstrated just how important it is to have data, there was a real feeling of urgency and that we could not wait another 12 months to get the programme right. That urgency led to the restart much quicker than perhaps people were ready for. There was also a feeling that people would be more accepting, and there would be fewer critics because people understood that data was important. Clearly, that was wrong. It has been proved to be a real issue.

The other piece is that the level of mistrust at the moment is not just about use of GP data. It is much wider distrust in the system as a whole at the moment. People still trust the NHS, but when people talk about a sinking feeling and us being back to where we were in care.data days, I think we are in a slightly different place. It is a place that I am much more worried about because for the first time people are aware that this data collection is happening, and they really get the benefits. They are saying, "We absolutely support the use of NHS patient data for research and planning to help the NHS. We get that. We support that. But we just don't believe that that is the only use of the data that is going to happen." That is the piece that is really concerning; the safeguards are

not being talked about enough, so people do not trust what else the Government are going to do with the data.

Q23 **Dean Russell:** To clarify that—I am conscious that other Committee members want to come in, so I will be brief—would you say that the mistrust is misplaced?

Nicola Perrin: I would say that the mistrust of what NHS Digital has in the way of safeguards, particularly with the new announcements about TREs, is misplaced. I think the safeguards are there. Healthwatch England released its polling data today. That shows that people have a very high level of awareness of the data collection. They think they have a high understanding of what is involved. When they have been shown NHS Digital's information and video explaining what has happened, nearly half of them said that the stories they have heard in the media and what they are hearing from NHS Digital do not match up. I think there is a real problem with misinformation about what safeguards are there.

That is why a really on-the-front-foot comms information campaign, with full engagement, is going to be so important—

Chair: Thank you. I am going to move on because lots of people want to come in.

Q24 **Anum Qaisar-Javed:** A lot of the conversation and the terminology that is being used in today's meeting is at quite a high level. I think it is really important that we take a moment to remember that there are members of the public who are genuinely interested in this.

I will use myself as an example of someone who has been treated by the NHS. I am 28 years old and female. As you can tell by my accent, I am Scottish, but for this argument we will pretend that I live in England. My ethnicity is Pakistani. I was born with a ventricular septal defect, which as Professor Landray will know is a hole in the heart. Unfortunately for me, mine got bigger rather than getting smaller. I had to get open-heart surgery that was quite invasive at the age of four. A couple of months later, I had to get my appendix taken out.

This surgery, as I understand—Professor Landray will give more information—is now a lot less invasive and the recovery time is a lot less. It is not as significant as it was back in 1996. That is because we have shared data and research. There is no doubt that can be a positive thing. I am not going to take away from that at all.

However, what I want to concentrate on, looking at my historical medical data, is that if I did not opt out, and I was still involved in this and I lived in England, how much of that information could be sold off? Would I know how far my information went? Could I possibly leave some sections of it out? If I did not want it known that my appendix had been taken out, for example, could I stop that?

At the absolute heart of this, it is really important that we also remember that many people trust the NHS. I trust the NHS. It saved my life. Could I

trust, if I lived in England, that my personal matters and health records would not be sold off? I am not convinced. This is an open question to all three panel members.

Chair: Who would you like to answer first, Anum?

Anum Qaisar-Javed: I have no preference. Perhaps Professor Landray because he is at the top of my screen.

Professor Sir Martin Landray: There are two things, if you lived in England. The first is that you could decide that data would not leave your GP. What that actually means is that your data would continue to be held by one of the commercial system suppliers that are under contract to your GP. GP data in England is held, by and large, by two companies. One is called EMIS and the other is called TPP.

The type 1 opt-out is, essentially, "I don't want data to move from one part of the NHS to another part of the NHS," from EMIS under contract to the GP, to NHS Digital, a Government organisation that is part of NHS England. That is where the type 1 opt-out comes in.

That poses some difficulties to me. This is my personal view. There is then no information about you. If others like you did the same, there would be no information about people like you. Even for fairly routine planning-type things or audit-type things, such as variation in care and all those sorts of things, the NHS would be flying blind about you and potentially people like you. To my mind, that is a problem that needs to be explained. We are talking about the type 1 opt-out, which is, essentially, about moving data from one bit of the NHS to another bit of the NHS.

You could also apply a national data opt-out, which would then stop, as I understand it—there are others on the call later who can clarify—any of your data, whether it is from primary care, your GP, or from secondary care, the hospital, being used for research. That is all or nothing, in the sense that you could not say, "I don't want people to know about my appendix but I do want them to know about my heart." It applies across the board. For a variety of reasons, and we can discuss another time whether they are good or bad, there are these two different types of opt-out.

As I say, my personal view is that the type 1 opt-out is not particularly helpful. I do not think it really gives people what they want. The final part of the question—

Chair: Sorry, Professor Landray. I have to ask everyone to be brief with their answers because we have the second panel to come. Anum, do you have any further questions?

Q25 **Anum Qaisar-Javed:** I would be interested in hearing Phil Booth's views on this.

Phil Booth: We are talking about the consent side of things now. We say that every use should be consensual. I am afraid that people are arguing from positions of their own interest. We advocate for patients, and we believe, as I understand the Chair does, that every patient's choices should be respected. That is done by an imperfect mechanism that we have been advocating for years should be changed and improved. Type 1 is the only option that protects your GP record, which is the most comprehensive bit of medical history that you have. It is much richer and full of stuff than your hospital history. The national data opt-out—the actual releases of data—is not respected about 80% of the time because of exceptions and because currently NHS Digital, like other parts of the NHS, does not treat pseudonymised data as personal data.

We can talk in theory, but please let's look at the facts. There are things that have been said to be done, such as the safeguards that I agree are there, but the learning that has not been done is that those safeguards have to mean what they say. Until they do, we cannot communicate them to patients. It is not just the words you say; it is that they are true.

Chair: Thank you. I am going to move on, Anum, but we will come back to you in the next panel. Rosie, it's your turn now.

Rosie Cooper: Thank you very much, Chair. Once again, I go last on a panel, which means—

Chair: You are not last. Someone is coming after you, Rosie, so go ahead.

Q26 **Rosie Cooper:** That is even better. I would like to address a number of points. Phil, the Minister and other people are going to come on afterwards and you will not have a chance to challenge what they are saying. Lord Bethell has turned his camera off and was looking distinctly cross before he disappeared.

Phil, you talked about theysolditanyway.com. Can you explain that a little bit more? How long, to your knowledge, has that been going on? Do you know of any work done to deal directly with companies that misuse that information? Did the NHS, NHSX or even the data guardian do anything or speak out about the practices of the companies that you refer to when you talk about theysolditanyway.com? Has anyone done anything? After all, quite clearly, as Sir Martin said, the HIV community chose to have their information regarded for research. Quite right. They chose and they had a say. Where do patients get a say? Where do I get a say? Who is going to protect us if this lot aren't doing it? Martin, over to you.

Professor Sir Martin Landray: Is that to me or to Phil? I wasn't sure.

Rosie Cooper: My apologies—to Phil first.

Phil Booth: With regard specifically to theysolditanyway.com, they are just re-presenting NHS Digital's own data release registers, which it had to do after, and as a consequence of, the Partridge review in the care.data outcry of 2014. In terms of what it could do, the contract says

it could get people to delete data. Obviously, that is after something has happened, if they are caught doing something bad, but you will see that people have breached agreements and breached the law, and are still receiving data, because certain regulations that we were promised back in 2014 to do with the processes of dissemination never surfaced. NHS Digital does not feel that it has the statutory backing to be able to ban people who break the rules from getting more data. If they do it, because some of these are commercial outfits, I guess they fear they are going to be sued, or something.

We are not there yet. There are some good signs. We are not quite sailing in the right direction. We don't have everything going along as usual, as it should be. We are thinking about turning the boat and we have done some moves. There are decisions like trusted research environment, fixing the opt-outs and writing to every patient. If we are not committed to that, these debates will continue to go on in public until people know that they will be written to when everything is fixed. All these doubts will still rattle around. There is no point trying to say, "Well, you should think that marketing is this way, you should think that anonymous data is that and you should think this and that, and it is all good for you." People have choices and they need to be able to make a properly informed choice. That means giving them the information they need in the language and at the level they need.

Q27 **Rosie Cooper:** Thank you. Martin, do you want to comment on that?

Professor Sir Martin Landray: I have little to say. Given that we are short of time, what I would say is that we have to be very careful about the use of "sell". The idea that data moves from NHS Digital to an organisation, and that that organisation does not pay for the effort of doing so, creating that, maintaining it and so on seems odd to me. I think we have to be very careful about "sell" as opposed to "paying a reasonable payment for the work that is required to do that, and to develop, maintain and grow the service". I think that is something that you might want to explore later on.

The second thing is that there are some real discrepancies. As was highlighted at the beginning, secondary care data—hospital data—has been shared in these ways with the safeguards in place for decades, and huge value has come from that. Yes, primary care data is important, and it is actually on the same people, but somehow we have a differential level of concern, at the same time as companies like EMIS are distributing data separately. We are happy for EMIS to distribute data or for TPP to distribute data, but we are not happy for this.

Phil Booth: No, we are not. They are data processors under the GP contract. If you can point to EMIS or TPP having distributed data that they were not directed to do, then please do so.

Professor Sir Martin Landray: It is under the same arrangement that NHS England and NHS Digital distribute data. The question is not a legal

one. The question is from a patient perspective. Are GP data reaching others? The answer is that there are ways of having GP data that are quite outside this. We should really focus on this discrepancy between psychosis information or hysterectomy information from secondary care being available, but information on heavy periods or anxiety disorders from primary care not being available. We just need to take things in the round.

Chair: Thank you. Do you have a final question, Rosie?

Q28 **Rosie Cooper:** I have a final question for Phil. Sir Martin has said it is not the same as care.data, but to me, having seen what has gone round the clock once before, it shares all the arrogant qualities of care.data. We talk about it being the patient's data, and it is the patient's data. They get the choice and the say. This, "We're doctors and we're doing it for the good of you all so you'll take what you're given" just cannot be. The use in this conversation of words like "slippery" is really scary.

If this is all so clear and beneficial, in your opinion, Phil, why did the Government try to get it done before anyone noticed? There is a lot of money to be made, and the public do not trust or understand what is going on.

Phil Booth: Again, I cannot talk to someone else's motivation. I think Nicola may have described a process that went on; that they thought they desperately needed this. What Professor Landray has referred to is the GPES Data for Pandemic Planning and Research. That is the GP data that has been extracted for the research that went into the RECOVERY trial and other research. That has been fine.

If we are talking about pandemic uses, which I think is the stuff that everyone completely agrees about, that is already happening and has been happening for a year. Why they decided to go for this generalised data grab, as I have characterised it, is that if you decide to do something in the middle of a pandemic, when you have not told GPs what you are doing in plenty of time so that they are ready for any patient reaction, and you initially put a six-week deadline to inform 50 million people by putting some stuff on Twitter and a website, it has all the characteristics of a desperate attempt to get the data and then apologise afterwards.

That is not the way it is going to work, and it is certainly not going to work ultimately for research either. Research has something far more than just legal compliance. Research and medicine have ethics. If you go about this in a way that is not ethical, it is not going to be as useful as people think.

Chair: Thank you. I am sorry to move everyone along. I want to make sure that we have time for the Minister at the end. Paul Bristow wants to come in on this panel before we move to the next one.

Q29 **Paul Bristow:** Thank you, Chair. I refer the Committee to my entry in

the register of Members' interests. Like Anum, I would like to talk about my own personal experience, very briefly.

A few years ago, I developed myocarditis. I was treated at the Royal Brompton. I did not opt in, but I was thrilled to have my patient data used in order to track what might lead to heart failure and use my personal experience and my personal data to help people in the future. That is something that I am incredibly pleased with.

The letter released by Jo Churchill yesterday talked about the ability to opt out for patients and strengthening patient opt-out. It talked about the trusted research environment. It talked about a campaign for patient awareness. These are all concerns that have been raised by Phil today. Phil, you must be thrilled.

Phil Booth: We are very supportive of the trusted research environment, as we have said. We have been saying for years that they should be sorting out consent. It is fine for people, individual members of this Committee even, who have the right information to make the choice that is right for them. We believe that should be true for all of your constituents and everyone in the country as well. That is why I have focused in my answers this morning on the poor communications.

Q30 **Paul Bristow:** So you are happy with that letter from the Minister.

Phil Booth: The letter from the Minister makes some good promises. Trusted research environments sort out a lot of problems, as I said. But to say that they have just done another communications campaign misses the point.

Q31 **Paul Bristow:** Okay, so you are thrilled with that letter. Hopefully, that reassures you on some of the points you have made today. Are there any circumstances in which you would opt in or recommend to your family or friends that they should opt in?

Phil Booth: When every use of patients' data is consensual, safe and transparent in the way that we have laid out, I would be perfectly happy to do that. As a younger man, I volunteered for drug clinical trials. I am one of the signatories to the AllTrials campaign. I want there to be transparency about registration of clinical trials. I am not hostile to research. I am a great fan of research, as is my colleague in medConfidential. It just has to be done right.

Q32 **Paul Bristow:** Good. You presumably have no objection to the use of HES data, national registries or bodies such as NICE using secondary patient care data in the way they do at the moment.

Phil Booth: We do. That is why we want HES data to be in the trusted research environment as well. We are talking about a process and a whole bunch of different datasets. We have advocated for patients' data across the piece. You cannot just do it for one part. You need to do it for all. We need to find mechanisms like a trusted research environment. It

would be ridiculous just to have GP data in there and not HES, or mental health data or any other dataset that is currently going on.

Similarly with the opt-out, as we have pointed out, it is not working for people.

Q33 **Paul Bristow:** Just to be clear, you want safeguards in place when it comes to primary care data, but you object to the use of HES data and the use of national registers in the way that they are used right now, in order to look at patient outcomes over a period of time, how technology and procedures have changed and post-market surveillance of the use of medical devices. You object to that as it stands.

Phil Booth: Not at all. It appears that you have not understood my argument. I am in support of all the positive research uses—

Q34 **Paul Bristow:** But you think that more safeguards should be—

Phil Booth: There need be no conflict. That is not the same as objecting to the way that things are being done and saying that there should be another way, or that they could be improved. We are moving towards that with TRE but, as I said, it would be ridiculous to simply put that level of protection around GP data and not to offer the same level of protections—because if they work, they work—to all the other data as well.

Yes, I object, and I advocate on behalf of patients who object to having copies of their individual-level hospital data, as we say, passed to third parties, including commercial organisations, for the payment of money and to service other customers that are not published. You cannot know who has your data. That is not a situation that can continue—

Chair: Thank you very much. We are going to have to move on to the next panel. We have had a very lively debate. I want to say a big thank you to Phil Booth, Dr Perrin and Professor Landray for setting the scene in the lively debate that we have had in the first panel. Thank you very much for joining us. It is much appreciated.

Examination of witnesses

Witnesses: Lord Bethell, Simon Madden, Simon Bolton and Dr Byrne.

Q35 **Chair:** I now welcome the Minister for Innovation, Lord Bethell; the director of policy and strategy at NHSX, Simon Madden; the interim chief executive of NHS Digital, Simon Bolton; and the National Data Guardian, Dr Nicola Byrne. Thank you all very much for joining us. It is much appreciated.

Lord Bethell, perhaps you can bring us up to date on what the programme is intended to do. Very specifically, could you address some of the criticisms that were made by Phil Booth earlier, because I think that is what is on people's minds? In particular, there is the suggestion that pseudonymised data is actually traceable back to individuals. That

was a very clear concern. Secondly, even though doing that tracing back to individuals is illegal, he said that there were multiple breaches of, for example, the level 1 opt-out by the NHS and nothing happens. Thirdly, he said that the communication has not been good and there need to be letters to individual patients. Those were some of the objections, and there may be others. Over to you, Lord Bethell.

Lord Bethell: Chair, thank you very much indeed. I would very much like to take on some of the points made. I might turn to Simon Bolton to talk about pseudonymised data and the breaches. We have been in a dialogue while the session has been going on, and I think he has some very clear points that it would be worth him, as the CTO, making better than I can.

If you don't mind, I will very briefly summarise where I think we are in the programme. Professor Landray and Nicola spoke very clearly about how data is key to unlocking some of the medical advances we have the benefit of at the moment. It is just staggering to see the phenomenal advances we have at the moment, where the basic science that we have invested in for a long time is turning into new-step therapeutics at a staggering rate. I am unapologetically evangelical about trying to mobilise data to save lives in that respect.

We know—Chair, I know you know this—that the system we have at the moment, GPES, is clunky and cumbersome and there have been questions about its security. We think that updating it is very important to drive medical innovation. I think most people recognise that need.

Professor Landray spoke about how, in the pandemic, data has been used to run the NHS efficiently—on the shielding list and on the vaccines prioritisation list. It is not just clinical research; it is for essential planning that we need to mobilise the GP data in this way. We need to apply modern approaches so that everyone gets the service they need. That is why we have been very focused on getting the programme done.

I completely recognise that this is a very sensitive project on privacy, on security and on whether we are taking the right approach at all. It is very important, and I think the point made earlier is right; everyone has to feel fully involved, and that includes patients and GPs. We absolutely have to take people with us.

I also recognise that this has not happened in the past. It is frustrating that we got off on the wrong foot this time. We have delayed the programme as a response to that. That is partly because the pandemic has disrupted normal operations, and there has been a degree of stop and start. With the limited resources at our disposal, we have tried our best to do things well, but there have been difficult circumstances.

I will level with the Committee. It is also because this project has grown from a reboot of the old system into something more ambitious. In fact, it is much bigger than just a reboot. We need to step up to the challenge

of engaging GPs and patients more emphatically. I completely recognise that. That is why we paused. That is why the team have engaged heavily with the key stakeholders in the last month to listen to their concerns, including important operational practicalities such as the resources needed to catch up on opt-outs; we have not talked about that, but it is a very important feature. That is why Jo Churchill has written to the GPs to make it clear that, rather than deadlines, we are committed to mutually agreed tests on opt-outs, on communications and on trusted research environments before we commit ourselves to the new system.

I think with the letter, and with the new framework in place, we have a clear agreement on how to proceed. There is a weekly meeting scheduled, chaired by me, to ensure that we have a strategic focus. There is a clear organogram and an SRO driving the operations. From my conversations, I feel that we are in a much better place this week than we were before. That would be my summary of where we stand.

Q36 **Chair:** Lord Bethell, could you respond to the concerns that Phil Booth raised—either you or one of your colleagues? I think the pseudonymisation is a pretty key one. Is that one for Simon to come in on? Simon Bolton.

Simon Bolton: We have got too many Simons. Thank you, Chair. The pseudonymised point is really significant. The data is pseudonymised at source in the systems that Professor Landray mentioned that hold the GP data before it is sent to NHS Digital. Once within NHS Digital, in the current circumstances, if it is to be onwardly shared, it would be shared in a pseudonymised way but under strict controls. Any agreements to share data would go through what was acknowledged earlier by previous witnesses as an independent review from lay people, who would take a view as to whether the use of the data was appropriate. We would then reserve the right to audit whether that data had been used correctly. The pseudonymised data happens at source.

As was acknowledged by Phil Booth earlier, the onward sharing of pseudonymised data presents a risk. That is why we agree with Phil that doing that in a TRE would be a more appropriate way of allowing research on this data, which has immense opportunity for insight into healthcare.

Q37 **Chair:** Thank you. What about communications and Phil Booth's point about a letter being sent out to everyone? He said that it is not just enough to put stuff on a website and to tweet. Really, you need to send a letter out to patients. Lord Bethell?

Lord Bethell: We acknowledge, as I said in my notes, that this may originally have been envisaged as a system reboot but it is more than that. It is about a massive engagement of trust. I would very much like to start that with the healthcare system itself. We need to get through to the 6,500 GP practices and the 15,000 GPs, and also to the receptionists and to the data—

Q38 **Chair:** I am so sorry to cut in, Lord Bethell, but it is a really straightforward request. I tend to be on the side of being a data champion, but it does seem to me a reasonable thing that, if there is to be a very big change, you should write to NHS patients to let them know. Would you consider that?

Lord Bethell: I am sorry for taking too long, but if I have to be brief I think there is a sequence. I think you have to start with the healthcare system—

Q39 **Chair:** So you are willing to commit to that in due course.

Lord Bethell: Exactly.

Q40 **Chair:** Yes. Could you write to us and let us know when you are prepared to commit to it? I think "in due course" could—

Lord Bethell: Chair, I am not committing to it.

Q41 **Chair:** You are not committing to it? Why are you not committing to writing to NHS patients?

Lord Bethell: No; I am not.

Q42 **Chair:** Why not?

Lord Bethell: Because we have to start in sequence with the healthcare system itself. We have to talk to GPs in the healthcare system and get their engagement on the subject. Once we understand—

Q43 **Chair:** Sorry. I understand that. Surely NHS patients are part of the sequence. You're saying that they are, so why aren't you prepared to commit to writing to NHS patients at any stage?

Lord Bethell: I am not ruling it out. I am saying that I am not prepared to commit to it today. I want to understand what we need to do to actually win patients. Phil may have a view that a letter is the thing that is going to make the difference. I am not persuaded of that.

Q44 **Chair:** I am sure that others will have comments on that. I want to bring in your colleagues on the panel before I move over to my colleagues.

Dr Byrne, first of all, congratulations on your appointment as the National Data Guardian. It is a very important role. You are obviously filling big shoes, but we are delighted that you are doing the role.

What is your perspective? Your job is to champion the rights of NHS patients to make sure that their data is being used responsibly. What is your view of the concerns that were raised by Phil Booth and what Professor Landray and Dr Perrin said in response?

Dr Byrne: Thank you for the opportunity to come today. I should probably add that, in addition to being National Data Guardian, I am a practising clinician myself. I am a consultant psychiatrist in adult mental health in south London.

I recognise a lot of the concerns that Phil Booth mentioned. I think we have seen in the public discourse that people have very significant concerns about issues such as commercialisation, and the question about who might profit from data. The discourse around data itself has moved on in recent years, and perhaps the system has not caught up with that in terms of public awareness of things like cybersecurity and privacy concerns. If we say that data is pseudonymised, what does that mean? What are the safeguards around that?

I recognise those concerns, and I think the responses to them were helpful. I agree with what Nicola Perrin said; there is a real need to be clear in communication with people and address those concerns head-on. That is certainly advice that I have been giving the Government and NHSX. I think there may have been a perception in the launch of this programme—the soft, quiet launch of it without much in terms of communication and engagement—that the argument for the uses of data was made with the public through the Covid experience, and that the benefit of data use would be self-evident.

While I think that the public are now much more data mature, and many people get the benefits of research and system planning, the fact is that many people have very significant legitimate concerns that need to be addressed head-on in a very honest and transparent way. For all of us as human beings, concerns and risks often outweigh benefits. The key thing now is to move the dialogue on into very clear and open discussion of the benefits, but engaging with the risks and concerns head-on.

Q45 **Chair:** Thank you. I will bring in Simon Madden for any comments he has on the debate so far, and then I will bring in my colleague Luke Evans.

Simon Madden: Thank you, Chair. It may be worth briefly spelling out the plan for our communications and engagement approach. The key word, as well as communications, is engagement. This is not just about us being on transmit; it is also about us listening.

There are four key phases to the approach that we want to take. The first is a listening phase, where we want to listen to stakeholders and gather their views on how best to communicate with the profession and patients. That is just to underscore the fact that we need to get this right, and to give them the opportunity to inform the development of the communications engagement approach and the development of the programme.

Next, there is a consultation phase when we will have a series of events, where we can explain the programme and listen and capture feedback, and help to co-design the work. There has been some excellent work done in various parts of the country on public deliberation events. There was one in recent years in London called "One London", which was a series of events bringing members of the public and the profession together in a workshop approach. Those types of approaches really reap benefits. I think we are currently looking at how we might be able to

scale up that kind of successful model into a national approach that would cover all of the country.

There is then a demonstration phase, where we have to show that the feedback is being used, and that we are actively using it in the development of the programme and shaping the communications. Finally, there is delivery of the whole campaign. That is about the nitty-gritty of the products and disseminating the information that the public need, to be able to understand better our approach.

I said I would be brief, but perhaps I have spoken for too long. Those are, essentially, the four key stages.

Chair: Thank you very much indeed. Let me bring in Dr Luke Evans.

Q46 **Dr Evans:** I have a set of questions, leading on from that, about communication. First of all, Dr Byrne, you are a clinician, as am I. What would you say to a patient sat in front of you, to explain what is going on?

Dr Byrne: I think the first thing I would do is listen to what their understanding was currently and any concerns that they might be having, just to frame the conversation so that I knew where we were starting from.

I would want to go through what the purpose and the ambitions of the programme were, which I support. As we have heard this morning—

Q47 **Dr Evans:** There are many people in the public. Give me your spiel—what you would say. I am sat in front of you and I say, "Doc, I don't know about this data thing. What should I do? Should I opt in or should I opt out?" What would you say? What is the answer to that question?

Dr Byrne: Perhaps NHS Digital would speak to this better than me, as I am not responsible for the programme, but I would say that NHS Digital is trying to improve the centralisation and collection of data to improve health research and system planning, and there are ways it is doing that which improve the current technology it is done with, as it is now getting outdated and needs replacement. If they were interested in finding out more, I would encourage them to do so, including, if they had any concerns, being aware of the choices they had about opting out for data use as well. It is very key to hold that choice hand in hand with what the programme is doing at the same time.

Q48 **Dr Evans:** That is really helpful; thank you. That leads me on to the next stage of my question, which is to Simon Madden. We are going for GPD. I am a GP by trade. I know that it has SNOMED for joining it all together. We are talking about probably EMIS or SystmOne. What data are you guys collecting on that basis? Are we talking about anonymised blood pressures, diagnoses, age-related demographics and that is it, bulked off? Are we talking more about the treatment plan all the way through, so you can say, "Well, no, Mr Smith—although now anonymised—has this

whole longevity of his AF into his blood pressure problem, into his stroke and his medication line he is on"? Simon Madden, can you package that up for me and tell me what it looks like?

Simon Madden: Simon Bolton may want to come in on this as the technical expert, but I will do my best. My understanding is that NHS Digital, in the collection, will collect data on diagnoses, symptoms, observations, test results, medications, allergies, immunisations, referrals and appointments. Crucially, what it will not collect is the entire GP record. It will not collect the patient name and address. It will not collect the written notes of any interactions between patients and clinicians—the so-called "free text" notes. It will not collect images, letters, videos or documents. It will not collect medicines, appointments or referral data that are over 10 years old. It will not collect legally restricted data such as IVF treatment or issues relating to gender reassignment. I hope that is technically correct.

Q49 **Dr Evans:** That is very helpful. Simon Bolton, do you want to come in and answer if there is anything missed, or to add any flesh to that bone?

Simon Bolton: No, because I think Simon Madden has got it right. The thing I would add, Dr Evans, is around informed consent. To your first question, we believe absolutely in informed consent. What we need to do is present the facts and allow people to take their own decision. It should be based on that rather than any recommendation from an individual.

Q50 **Dr Evans:** That is really helpful because from my perspective lots of doctors do presentations to other colleagues about a patient, for example, who is 25, and has had an operation. It is pseudonymised but we can understand so that we can learn. That is how medical journals go on. Am I correct in thinking that, in essence, this is what you are trying to do, but on a national basis?

Simon Bolton: Yes, I think that is a good description of it.

Q51 **Dr Evans:** That is very helpful; thank you very much. That leads me to Lord Bethell. In the first discussion, we heard all about the pros and cons, and this is a passionate subject. Is it simply the fact that this is a good project but the Government have got the communication wrong?

Lord Bethell: No, I don't think that. I think it is very difficult. Personal data is extremely sensitive. In all walks of life—finance, retail and the internet—we are all having to come to terms with how our data is treated. Health is special data. It is particularly personal, so of course it has to be treated with additional reverence.

I think we may have made a mistake in thinking of it as an IT project rather than as something that has to engage with GPs and the public. We are not really set up to do that in the NHS, so we are having to create infrastructure and techniques to do that. We are up for doing that, and that is what we are focused on delivering at the moment.

Q52 **Dr Evans:** That is really helpful. You are right that data is so important, with all the discussions around social media and collection of data. What goes on in America in particular is often cited as a concern. Does that mean that the Government are having a rethink about how to communicate this strategy due to the sensitivities, and making the distinction that the NHS is so different, both in the protection and understanding of the data, and because of the place it has in our heart, given the concern about privatisation versus nationalisation?

Lord Bethell: I think we were always cognisant of the sensitivity of the data. In fact, if you spend time with NHSD and NHSX, you can see they are incredibly protective of the data. We have security, protocols and legals all around the data.

I do not think there is any misunderstanding of how sensitive the data is. Where I think we have to be more ambitious is in the scale of the engagement. It is going to be a big project. Simon Madden ran through a very useful four-point plan that has dialogue at the centre of the engagement. Your idea of, "What's your pitch to someone?" is a helpful question to get the articulation, but that is not how we will win this argument. We will win by listening to people and then answering their questions. Doing that on a large scale, as we have done with the vaccine, is something that requires a big investment.

Q53 **Dr Evans:** That is really helpful. This is my final question for you. What has surprised you the most about this project?

Lord Bethell: To be honest, I have now become a massive evangelist for it. The beneficial impact, both in planning and in clinical research, really can be transformational. I am staggered that we are not doing it already. I have massively prioritised it in my own focus, as we are beginning to come out of the pandemic, and thinking about what I have to do in the year ahead. I think that the benefits for saving lives, for making the NHS more efficient and for advancing science are so profound.

Dr Evans: Thank you.

Q54 **Laura Trott:** Dr Byrne, I want to ask you a few questions because your role as the National Data Guardian is very important both in the design of the system and in reassuring the public in communication. You said at the outset, in answer to some questions from the Chair, that the advice you have been giving to the NHS has been around making sure that people's legitimate concerns are taken into account. Has your advice been listened to throughout this process?

Dr Byrne: Yes. I came into post in April, after the decision was made about how the programme was to be launched. Since that point, I have been in communication with NHSX and the Government about it, and flagged concerns about the timescale, and not only the communication, as Lord Bethell and Simon Madden have spoken to, but the question of actual engagement. That is not just with the public but with the professions themselves. It is very important that frontline GPs are

directly engaged, too. I think the timescale was ultimately unrealistic about the time in which those things could happen.

I am very pleased by the letter that went out yesterday from Jo Churchill. It is clear to me that my advice, and that of others, I should add, has been very carefully listened to. I am very pleased that they have announced a delay, and the reassurances in that letter are very important.

Q55 **Laura Trott:** The timescale and the engagement are your two concerns. You had no fundamental concerns with the design of the programme?

Dr Byrne: The solution now to move to more rights and a TRE structure is a good one. I think it addresses what lots of people have raised around privacy concerns and will provide a stronger safeguard than the infrastructure that was initially proposed.

The problem with the initial proposal was the lack of trust around it in NHS Digital, and lack of awareness of what the safeguards were. That said, I absolutely support the TRE solution. Now, certainly one thing I am advising, which has been taken on board, is that more time is needed, not only for the comms and engagement but for resourcing the infrastructure change that is, in itself, key as well.

Q56 **Laura Trott:** Having now satisfied yourself that you are happy with the system and the way it is being viewed, what do you see as your role in helping with outreach to patients and professionals, as you rightly highlight?

Dr Byrne: I was invited and am very pleased to join an independent advisory check and challenge group for the programme. Through that, we have spoken about the things I have said already, but I also want to advise on the opt-out process, looking at that and making it much simpler, so that it is a simpler process for people than currently happens. As part of this, it is really important to be pushing for the development of user-friendly data release registers. That is something my predecessor called for and it is in my priorities for next year, which will be coming out soon in my annual report.

Listening to the questions and the discourse this morning, I think there is a lot of confusion about how exactly this data has or has not been used. A lot of those questions could have been addressed if we already had user-friendly release registers that people could look at themselves and interrogate. NHS Digital already provides a use register, but it is fairly impenetrable. You would have to have a technical mind to work your way through it. All of us in the public should be able to access that and interrogate it ourselves to see if we are satisfied about how data is and is not being used.

Q57 **Laura Trott:** Do you think the user-friendly data release register needs to run concurrently with any resumption of the programme?

Dr Byrne: I know that work is already happening on that. I am very keen to support it in an advisory capacity. My predecessor called for a tool so that people could see how their own individual data had been used. I am very pleased to see that developing that is in the Government's data strategy. I think that is important too, alongside how all our data as a collective is being used. People are right to want to know, "Has my data been used too?" I think those two things can happen concurrently, and both are important.

Laura Trott: That is very helpful. Thank you, Dr Byrne.

Q58 **Dean Russell:** If I may, I will address my questions initially to Lord Bethell. I have a bit of a comment to begin with, if I may. Forgive the background noise; we have building behind us.

One of the things that seems really interesting to me is around data literacy. Over the past year, through the pandemic, people have got more used to data and looking at statistics and analysis, and understanding the role of an individual data point within the context of much more.

With that mindset, are there learnings we could take when we talk about the communication of this? People get data more, and they understand that they are sharing every day on social media and other points parts of their lives. Actually, the opportunity to save lives is perhaps the key point. I look at things like organ donation programmes and the communications around that. Should we be looking for the communications to be more about data donation—"You're not just saving one life but saving thousands by donating that data"—rather than it being a more technical conversation around opt-in and opt-out? I wonder what your thoughts are on that.

Lord Bethell: I completely endorse that. The organ donation example is a good one, and so is the vaccine. A lot of data was used in order to put together the prioritisation lists. My inbox is not full of complaints from people who got the vaccine wondering whether their data was used in delivering it. It really isn't.

I think that the public health literacy that you allude to is right. We wear masks not to protect ourselves but to protect our neighbours. For a lot of people, having the vaccine is not to protect themselves but to protect their loved ones and the community. There is a lot that people have done in the last year and a half that was, frankly, selfless and was to protect the community at large. That is what we are talking about here with the data. The individual benefits of sharing data are minimal. The communal benefits are enormous. People have twigged that that is what being part of a national health system is all about.

I agree that, at the same time, there is concern about social media and about other uses of data, and how the big tech companies use our data. That literacy is increasing, and the concerns are increasing at the same time. We have two, big related public learnings going on at the same

time. I am really hopeful that, as people think more about how their data is used, they will join the communal effort.

Q59 **Dean Russell:** Thank you. I have one other question, although it is more of a thought. One of the comments so far has been about anonymisation of data. Something I think I might have mentioned in the House previously was about the concept of creating something called placebo data, whereby you automatically introduce two or three points of data in every dataset that are slightly different, so that you could never actually trace it back fully to an individual. Is that something that has been discussed or thought about?

Lord Bethell: Could I pass that to Simon Bolton, who is more familiar with these matters?

Simon Bolton: I am not quite sure that I understood the question.

Q60 **Dean Russell:** Given some of the examples earlier, where somebody might be of a certain age, a certain ethnicity and with a certain number of children and so on, are there opportunities to help anonymise data a little bit more if you introduce two or three variants within everyone's dataset that are slightly different, as an individual set of data? They would not directly match that, but overall it would effectively become noise, so it would not affect the overall analysis, if that makes sense.

Simon Bolton: It is really important to understand how we are using pseudonymised data currently. If I talk to the risks that were identified earlier around the opportunity to market to people as a result of data sharing, we simply do not believe that is possible at the kind of scale that was implied. The way we manage this is that we pseudonymise the data. We then have a very rigorous process before we release that to third parties to do research on. We have a contract in place with those third parties detailing exactly what they are allowed to use that data for. It certainly would not include marketing. Then we do audits selectively, where we believe there could be risk, to ensure that they are not using the data inappropriately and are not doing things like marketing, which would be almost impossible to do because they would have to re-identify the data. They may be able to do that in one or two cases, but they certainly could not do it at scale on the datasets we have. Does that answer the question?

Q61 **Dean Russell:** Yes, I think so. It is more about how we make sure that it is not trackable back, and the concept of placebo data or something else that is introduced to ensure that was not possible.

Simon Bolton: Our commitment now is that, instead of sharing data externally, we will do all of that analysis within our own infrastructure, in a TRE. We will be in control of exactly where the data is. It could not be forwarded, as could currently happen.

Q62 **Anum Qaisar-Javed:** My questions are directly to Lord Bethell. They are nice and short, with, hopefully, nice short answers.

Lord Bethell, there is very little information on what impact this will have in Scotland. While we assume that there would be no impact because health is, of course, devolved, the former Health Secretary spoke of one NHS across the United Kingdom.

Can you clarify two matters for me, please? Will there be any data sharing from Scotland? Yes or no? When it comes to engaging with devolved Administrations and the Cabinet Secretary for Health in Scotland on the use of all this matter, will it be a case of consulting Humza or gaining consent from the Scottish Government and the NHS in Scotland?

Lord Bethell: On the second question, we have an enormous amount of engagement. In fact, I had our engagement committee last night with Humza chairing, and we are proceeding, particularly in areas of research where there are some blurred lines between what is national and what is devolved. On that, collaboration and engagement is absolutely the priority.

On whether data is actually being shared at the moment, I could not answer conclusively. I would like to ask Simon Madden, who may know the answer, or else I am afraid we will have to write to you. Simon Madden, do you know the answer to that?

Simon Madden: My understanding is that there would be no sharing with the Scottish NHS. There are some provisions where Scotland's NHS and public health infrastructure, for instance, share data on Covid-related issues for the response to the pandemic so that there is a UK-wide position, but for the purposes of this particular programme, as I understand it, there is no impact in Scotland.

Anum Qaisar-Javed: So I can leave this Committee meeting knowing that there will be no impact on Scotland. Thank you.

Chair: Is that you done, Anum?

Anum Qaisar-Javed: Yes. I said it would be nice and short.

Chair: Great. Thank you very much.

Q63 **Barbara Keeley:** I will address this first to Lord Bethell, although other panel members may have things to say. I want to reflect on what we have heard about the current situation. It is important to go back to what Jo Churchill said as Minister. She said that personal data belongs to the patient, when she was talking about this in the House, and they allow the NHS to use it for health and planning reasons.

It seems to me that we have ended up in a situation, again, where privacy, which is so important to patients, is pitted against data use. I reflect on why the Government and NHS Digital have made such a mess of communications on this that millions have now opted out—maybe 3 million.

To come to the current situation, the pause should help improve

communication. This is a direct question to you, Lord Bethell. Will you use that pause to send a letter to each patient explaining this system and its use, and giving information about how to opt out? I think you need to explain to patients, given the mess that there has been, and following on from the mess of care.data—a legacy you have to live with, I am afraid—how you have fixed what has gone wrong.

Lord Bethell: Barbara, I completely accept the challenge that we have to explain this to patients. I completely accept the challenge that these are complicated issues where a lot of people have grave concerns about security and privacy that we absolutely need to answer. The philosophical concept that the data belongs to the patient, which I believe emphatically, is not wholly borne out by either the technical or legal circumstances of the moment, and we are keen to try to resolve that.

On your specific question about the letter, I am just not persuaded that that is the correct way to engage with patients in the UK. We do not send letters very often. They can be a clunky form of communication. If it emerges that that is the right way to engage, I will be happy to do that, but I am not prepared to commit to that today.

Q64 **Barbara Keeley:** I would just remind you that the Prime Minister sent a letter to every household at the start of the first lockdown last year. It would not hurt to think through this issue. It is one of the most important things you could do. As to the communications that have been described to us—workshops in one area of the country and that type of thing—I know from all my attempts as an MP, and in other roles, to engage with the public, that what you are describing is engagement with a very small number of patients. I do not think it is good enough, so let us leave that there.

I want to talk about the national data opt-out. Will the Government put the national data opt-out on a statutory footing in the Health and Care Bill? Can you comment on the points that Phil Booth made that the national data opt-out is not respected 80% of the time? I guess that means in contracts that medConfidential has looked at. We have a national data opt-out. Will it be put on a statutory footing, and what will you do about the fact that it is not being respected?

Lord Bethell: On the letter, the PAC said that a letter was not good value for money and would not have the desired effect. We did a letter on care.data and it did not work out. I am sorry, Barbara, but I have to do things that work and engage people. I want to engage patients, but I will not be held to a specific form of communication over any other.

A statutory footing is something I am prepared to look at. It is not currently part of the plan. I think that the opt-out is very important. We have made it clear that we do not think it works as well as it should. We are doing a lot technically to sort out the different opt-outs and to catch up on the backlog. I think that patients deserve to have more transparency and visibility of how their data should be used. If that

requires a statutory footing for people to have confidence in it, I am prepared to look at it, but it is not on the schedule at the moment.

Q65 **Barbara Keeley:** Counter to what we have heard, it sounds as if it is being suggested that it is not possible for data to be used for marketing. While we have been in this Committee, I have looked up some of the people who have contracts currently to use our patient data. I will give you the names of two of them. One is Experian Marketing Services, which combines HES data—admitted patient care data—with other data from Ofcom and Police.uk websites and sells profiles on to public and commercial organisations. Harvey Walsh is another organisation that uses hospital episode data and then provides services to commercial organisations.

At the moment, with that being the case, we cannot say that we are not selling on patient data. We are selling on patient data. Firms like Experian are known to use data in those ways. We know what Experian does in terms of its business. Harvey Walsh is another of that type.

Lord Bethell: Barbara, if I can come back on that, it is not authorised in the agreements on the data to use the data for marketing. We have a comprehensive audit process around that. If you can point me to precise examples of where people have breached their agreements, I would be very happy to follow them up, but smearing the system with accusations like that does nothing to build trust in something that is extremely important.

Chair: Can we bring in Simon Bolton on that point? He had his hand up.

Q66 **Barbara Keeley:** Yes. Can I be very clear that criticising and challenging after the mess that we have seen over the last few months is not smearing? That is an unfortunate word for you to use in this Committee. We are trying to represent the interests of patients. It is not smearing to point out that currently their data is being sold on to Experian Marketing Services and firms like Harvey Walsh. There is a long list of other organisations, but they currently have a contract.

Lord Bethell: We do not sell the data. We ask for people to pay for the transaction costs—

Q67 **Barbara Keeley:** You are playing with words.

Lord Bethell: It is not our intention to sell it, and therefore to allude to selling it constantly is a misrepresentation.

Q68 **Chair:** Simon Bolton wants to make a contribution on that point.

Simon Bolton: Yes; if I can, please. I think it is really clear to understand that within NHSD you have to apply for the data. We go through a rigorous approach of understanding exactly what you are going to use the data for. We go through a lay body called IGARD, which includes GPs and others, who assess the reasons for that release of data. We only release it if it is for valid research.

Q69 **Barbara Keeley:** Does Experian Marketing do valid health research when it combines hospital episode data with Ofcom data and Police.uk data to sell profiles on to public and commercial organisations? Experian? We know what Experian does as a company.

Simon Bolton: We know what Experian does in general, but in terms of this specific data release I do not know. I will absolutely go and investigate that and find out what the release was for.

Q70 **Barbara Keeley:** And Harvey Walsh? There are others.

Simon Bolton: We would not, and do not, release data for anything other than genuine research opportunities.

Q71 **Chair:** Simon, would you write to the Committee with the answers on Experian and Harvey Walsh, so that we can understand?

Simon Bolton: I will absolutely do that. To re-emphasise the point that Lord Bethell made, because he is absolutely right, we do not release data to be sold for marketing.

Q72 **Barbara Keeley:** I do not accept that until we are clear about it. I would be grateful if you looked into it. There is another point to make. The word "anonymised" has been used in this Committee today, but it should be "pseudonymised". I split off in my mind the two risks that I think there are. There is snooping and hacking, the unlawful use of data to identify maybe celebrities from things like hospital records to try to find out what else is in their health profile, and there is the other use, the commercial use for which there is such a big demand. We should not mix them up. One is an unfortunate scourge on our society, and we have to live with it and make sure it is not there.

My final question is about the opt-out process, Lord Bethell. You have announced a review of the opt-out processes. Will there be a better opt-out process for families with children? Could you tell us how many opt-out forms are in the backlog at NHS Digital?

Lord Bethell: The opt-out for families is one I know about because I have had personal experience of it. If you don't mind, Barbara, I am going to pass that to either Simon Madden or Simon Bolton, who are technically more expert than I am.

Q73 **Chair:** Who wants to go first? Let's try Simon Madden.

Simon Madden: Sorry, Chair. There was a slight delay because we are in the same room sharing a microphone. We were just trying to make sure the logistics worked correctly.

We have committed to refining the opt-out. Part of that is that NHS Digital will have the ability to delete data if patients choose to opt out of sharing their data, even after their data has been uploaded. We want to introduce a more flexible system so that people can opt out and the data will be deleted, even if the data has already been uploaded. That ensures that patient choice is absolutely respected. It also means that we will be

able to have a more flexible system to enable people to opt back in as well, which is crucially important.

There is a current backlog of opt-outs, as I understand, with GPs, because type 1 opt-outs are administered in the GP practice. We have been discussing with the BMA and the RCGP how we might relieve them of that burden and find some way of centrally processing the type 1 opt-outs.

Q74 **Chair:** How many opt-outs have not been processed, Simon?

Simon Madden: My understanding is that we do not hold those figures centrally because they sit within GP practices.

Q75 **Chair:** I am sorry to interrupt, Barbara. I know these are your questions, but why can we not just decide whether to opt out or not on our NHS app? Isn't that the obvious place for anyone to be able to express their data preferences, Lord Bethell? It would seem the easiest way for a lot of people.

Lord Bethell: I think that is a good challenge. It is a very fair point. I do not have an answer for you, but let me look into it and I will get back to you.

Chair: Thank you. Barbara, over to you. Sorry I interrupted.

Barbara Keeley: Lord Bethell, you made the point that there was a letter sent for care.data. Both Rosie Cooper and I were around on this Committee when care.data was such an issue. There was not a letter sent to patients. There was a type of letter delivered to households; it was the sort of thing that comes with pizzas and it is about 30 seconds before people put them in the bin. I think a letter to patients, as we have advocated today, would be a very different thing. I will just leave that there. Thank you.

Q76 **Paul Bristow:** Lord Bethell, I asked a question in the last panel about the use of secondary care data as opposed to the use of primary care data. HES data is routinely published. The registries use patient data for post-market surveillance of health tech devices and med tech devices, and to try to improve clinical practice. NICE uses secondary care data all the time in the development of clinical guidance. Why do you think there is not the same concern around the use of secondary data as there appears to be over primary care data?

Lord Bethell: "I don't know" is the honest answer. There is something special about a patient's relationship with their GP. It is a very personal and intimate relationship and, therefore, the data that is held in that relationship is different in tone from what you might have in secondary care. There is a historical question, in that we have not used it before, so this is a change, whereas in secondary care it has been going on for decades and therefore does not feel different.

I basically agree with your premise. There is no difference. We are just seeking to update and modernise something that, in large parts of the healthcare system, has been commonplace for a very long time.

Q77 **Paul Bristow:** I was taken by one of the comments from Professor Martin Landray in the last session. He talked about the missed opportunities if we do not use primary care data on depression, sexually transmitted diseases and other things. Do you fear that if we continue to use words like commercial, marketing and sales—because language is really important—the opportunity around making advances in these areas could be lost?

Lord Bethell: I do. Simon Bolton spoke about the IGARD process, which is an incredibly rigorous process to ensure that clinical research is signed off to a very high standard. I really recommend that everyone has a look at their minutes because they are very thorough. The projects that they sanction are incredibly valid and worthwhile.

What shocks me about IGARD is what a small number of projects have been completed. The technical system and the administrative processes are so onerous and long-winded that we are simply not keeping up with the demands of the scientific community for this valuable data. I am worried that it is becoming a drag anchor on progress. That is where my appetite for trying to change this comes from. That is why I am passionate to try to get right the relationship between patients, GPs and their data.

Q78 **Paul Bristow:** I have one last question. Feel free to bring in some of the other witnesses on this if necessary. It is just a hypothetical question.

In the NHS, because we have a single-payer system and a single-provider system, the use of data could be transformative for the UK as a destination for clinical trials and scientific research. How much does that play into your thinking when it comes to using patient data in the way you propose?

Lord Bethell: Massively. We put the use of data and clinical research on the agenda for the G7 health summit. It was an agenda point with the other Health Secretaries of State from the developed nations. It was really striking how they all saw and were very envious of our opportunity in the UK for using the NHS as a platform for clinical research and driving innovation in therapeutics and devices. They could all see that the UK was fantastically placed for that. We desperately need to take advantage of that opportunity.

Paul Bristow: Thank you.

Q79 **Rosie Cooper:** Lord Bethell, you have previously said that opt-outs should be covered more. How many thousands of families are in the queue for NHS Digital to have their forms processed? I understand that is the only way to opt out if you have children. I get the difference between type 1 and type 2; none the less, how many are in the queue that you

can have access to, which is type 2?

Lord Bethell: Rosie, I will ask Simon Bolton to give you the precise answer. It is slightly mindboggling that this is the case, but because we cannot see into the GP data, we do not know how many are either opting out on a daily basis or are waiting to be opted out. That is not a great circumstance for a modern data system. It is one that I regret and would like to address. Simon Bolton, can you help at all?

Simon Bolton: Yes, I can help with the backlog for national opt-outs. A few days ago—I am afraid the data is a few days old—there were about 4,000 in the backlog. To be clear, you can perform the national opt-out through the app, for which we now have 10 million users. In fact, you can opt back in through the app as well.

Q80 **Rosie Cooper:** Thank you. I will come back to Lord Bethell. You suggested in your earlier evidence that opt-out has not worked as well as it should. I am really concerned now that it might mean that previous opt-outs are not being respected. How can a patient check that their own opt-out has been respected, and is functioning and works?

Lord Bethell: Thank you very much for the chance to clarify what I said. I did not mean that if you had opted out your opt-out was not working or was not effective. I absolutely did not mean that. What I meant was that the customer experience is not as clear and transparent as it would be for a modern organisation.

It is not a direct comparison, but if you are a subscriber to Tesco or to the BBC, you have very clear and transparent processes around opting out and seeing your data. I do not think we are quite as good as we should be on that. I would like a trusted relationship between the patient and the NHS to be based on transparency and ease of execution. We need to improve on that. That is what I meant.

Q81 **Rosie Cooper:** You are saying that any opt-out is being respected right now.

Lord Bethell: Yes, absolutely. I am happy to clarify that with the team, just to check that I have not overlooked anything, and get back to you. That is certainly what I meant, yes.

Q82 **Rosie Cooper:** My next question is to Dr Nicola Byrne. I was a massive supporter of Dame Fiona. I thought she did an absolutely wonderful job in coming in and rescuing the mess that was care.data. Do you feel you could have been more publicly proactive in this current mess? The Government would already have implemented this data grab but for the work of campaigners such as medConfidential. Genuinely, what have you done privately and publicly to defend patients? The second part is, do you think that every patient, as it is their data, should be consulted? If so, how?

Dr Byrne: To take your first question about what I have done personally, I have—

Rosie Cooper: And publicly.

Dr Byrne: Put it this way, I made a decision when new in post that at that point in time the most effective and impactful thing I could do was to provide private advice to achieve the objective of the programme being paused, and that is what happened. I did not think it would be useful at that point to make a public statement in that regard if I was giving the advice in private. Meanwhile, other organisations were speaking publicly, very helpfully, I think, and raising very legitimate concerns.

As I hope I made clear earlier, I am satisfied that my advice, not only on the pause but on some of the specifics that the letter that came out yesterday to GPs addresses, has been listened to and taken very seriously.

Please remind me of the second part of your question in terms of patient control and choices.

Q83 **Rosie Cooper:** How should they be consulted?

Dr Byrne: I am very pleased to say that there is now an independent editorial board advising the Government on this programme. A member of my panel, with an extensive background in comms, represents me on that group. They now meet very regularly and are looking very carefully at the entirety of the comms and engagement programme. A lot of attention has been focused on a letter. I think that certainly should be considered.

One thing I am very concerned about, and have spoken privately and publicly about, is the importance of putting this programme in a wider context of data use. It has been interesting that even in discussing this today, we have often veered on to hearing about secondary care data—the HES data—and cancer registers, for example.

A concern that I have is that in recent months many members of the public might have had the impression that currently data is not being used from primary care, when it is, quite extensively and with, I would argue, less secure safeguards around its use; and they might think that GP collection is the only thing happening with data, and that is not the case. This is advice I am giving on the data strategy as well: it is really important to put this programme in the wider context of data use and to make clear to people all the different flows that are happening, not only from primary care but from secondary care and social care. Let's not forget that in the picture. It is really important to bring these things together so that there is a coherent narrative, rather than just coming away with the impression, as I think many people might, that it is all just about primary care.

Q84 **Rosie Cooper:** Thank you. I would ask you to consider whether you are

the Government's data guardian or the people's data guardian, and whether your role going forward will be substantially damaged if people just think you are there as the Government's data guardian and not as a public voice speaking for them.

Might I make the point to Simon that the public will be rather discomfited that you, as the chief exec of NHSX, do not know what Experian is using the data for and what the contract says? I do not know of any medical use Experian would use, other than a data marketing type thing. I accept that you are going to write to us, but the public and somebody like me will be aghast that you, as chief exec, have to write to me and do not know.

Dr Byrne: Before we go to Simon, can I pick up on a point you made there? I think you are right. It is very important that the public know I am their independent champion and am here to protect the patient interest—the public interest. That is absolutely right.

To add to that, perhaps I could reflect on an earlier question you asked. Before I was appointed, I was lucky enough to speak to Dame Fiona Caldicott on several occasions. Of course, in that context she gave me some very helpful advice. That included the importance of navigating the decisions around what one makes private and public. I absolutely agree that I am here to be independent. In terms of this morning's Committee, I was rather perturbed to be listed, if you like, in the Government session when I am not a member of the Government. That distinction is very important for my role, and for the public to be able to trust me in the role.

Q85 **Chair:** You are here to give challenge to the Government, Nicola. That is why you are here.

Dr Byrne: Exactly. That's what I am here to do.

Q86 **Chair:** Rosie, we will have to wrap up soon. I will let Simon Bolton come in because he has a brief comment.

Simon Bolton: I feel that I need to respond to Rosie. First of all, I am the chief executive of NHS Digital and not NHSX, just to be clear about that.

Rosie Cooper: Apologies.

Simon Bolton: We have a lot of data-sharing agreements. I think it is reasonable not to have the detail of all of them. I know, for example, that Experian has public sector contracts with CCGs. It does a lot of work in this space. For example, it did some work with the lung disease charities around lung disease in the Isle of Wight. It does work in this space, but I do not have details of the exact data-sharing agreements that we have in place for Experian.

Rosie Cooper: Thank you.

Chair: Thank you very much. We have come to the end of the session. I

just want to make a quick final comment to Lord Bethell, if I may, before we wrap up. We very much appreciate your time this morning. I know that you are going to write back to us with various bits of information that you did not have to hand. Could I ask you to add to that?

The one thing where there is the most misunderstanding is the question of what the data is going to be used for. In the first session, there was something Nicola Perrin—a supporter of data being used for research purposes—said that was of the biggest concern to me. She said that, although we believe this data is being used for research, a lot of members of the public think it is being used for marketing purposes. What would be very helpful is if you could specify what the data is being used for, and also specify what the checks are and whether there is actually any independence in the checking process, so that the public can really know that the data is only being used as intended.

You are also going to let us know what the Experian contract is about. There was another contract that Barbara Keeley mentioned. Clarification in that area would be extremely helpful.

Thank you all very much for your time, the two Simons, Lord Bethell and Dr Byrne. We really appreciate your helping us to shed some light on this very important issue this morning.