SCOTTISH HOSPITALS INQUIRY

Hearings Commencing
19 August 2025

Thursday, 21 August 2025

CONTENTS

THE CHAIR: Good morning to everyone in the room, and I obviously include Dr Drumright in that. Good morning, Dr Drumright.

Dr Lydia Drumright
Affirmed

Now, Dr Drumright, as you will understand, you're about to be asked questions by Mr Mackintosh, who I think you've previously met. You're scheduled for the whole of today. I anticipate that your evidence probably will take the whole of the day. Our timetable is that we usually take a coffee break at about half past 11, we'll break for lunch at 1, take an hour, and then sit again at 2 with a view to completing your evidence about 4 o'clock, but that may or may not be precisely our end time. However, if at any time you want to take a break for whatever reason, just give me an indication, so please feel that you're kind of in control of that side of things. Now, Mr Mackintosh.

MR MACKINTOSH: Thank you, my Lord.

Questioned by Mr Mackintosh

Q Dr Drumright, I wonder if I can take your full name.

A Lydia Nicole Drumright.

Q Now, Dr Drumright, you are author of what we have come to know as the HAD report and, before I start, I want to understand a little bit more about your background. So I wonder if we can go to your CV, which appears in bundle 44, volume 1, page 196. It should appear on the screen in front of you. I'm obviously not going to go through your CV entirely, but what I wanted to understand is where are you currently based?

A I'm at the University of Washington.

Q In----

A In Seattle.

Q In Seattle. Now, did we fly you over here from Seattle, or were you perhaps in the UK anyway?

A No, you did not. I was here visiting family, so I'm a British and US citizen and I have family here as well as in the US.

Q Well, thank you for coming up to Edinburgh. What would you describe your specialism as?

A I'm an epidemiologist. That's my primary training and specialist, and I am also a health informatician.

Q Now, we've heard other people describe themselves as health data scientists. How does that relate to being a health informatician?

A So-- That's an interesting question, thank you. Data science

became a new catch-all term for people who were not necessarily what we call card-carrying statisticians, so they did not have a degree in statistics, but did data science. So I could call myself a data scientist; I could call myself a number of things. A health informatician is slightly different in that we build electronic tools, work with electronic health data, and that's the sort of stuff we do to support healthcare.

Q And does your professional experience include working inside the NHS in England on health data system?

A It does. So I used to-- When I was at the University of Cambridge, I led Cambridge Clinical Informatics, which was established to analyse our data scientifically from the Epic system when we put it in years back.

Q And what would you say is your principal research interest at the moment?

A So, my principal research interests are split in a few different areas. I work on epidemiological problems, including-- I do work with Dr Samir Agrawal on bloodstream infections and fungal infections in haematology patients, as well as working on chronic disease development in people living with HIV, and also health informatics tools.

Q Your original degree, I think, was biochemistry and cell biology.

A That's correct.

Q To what extent would you consider yourself a biochemist?

A I would consider myself a biologist over a biochemist. So, my background focused a lot more on infectious diseases when I was doing my biochemistry degree.

Q And obviously we've had a lot of evidence from a lot of microbiologists, whether academic or in practice in the health system. To what extent do you feel comfortable getting into the territory of microbiological questions rather than epidemiological questions?

A It depends on how far you want to go with those. If you want to understand extensive detail about bacteria, I would defer you to my colleague, Professor Hawkey.

Q Yes, and so it's a joint report, and in a moment, I'll come to how we might divide the work up, since you're the first of the three. But what's the extent of your experience in infection prevention and control in hospitals?

A So, when I was in the UK, my first post in academia was at Imperial, and I did a lot of work with the infection control team. So, Professor Alison Holmes runs that team still today, unless she's moved recently, and I worked quite closely with them but on research questions.

Q And so when we-- We've spent a lot of time in this Inquiry discussing how one investigates a potential healthcare-associated infection through PAGs and IMTs and all these sort of practical day-to-day processes. What's the extent of your knowledge or experience over that side of IPC?

A Yeah, I have not spent any significant amount of time sitting in those processes. I'm aware that they exist, I'm aware that they go on, I know some of the details, but I would not consider myself an expert in those processes.

Q Has your work as a health informaticist involved feeding epidemiological information to people who have been doing IPC work in hospitals, whether here or in America?

A So, that is what we were working toward at Imperial when I was there before I moved, as well as the University of Cambridge. But again, I would say the clinicians would tell us what their problems were, and we were there really to be able to tell them how we might process data and present it to them-- is the extent of what I would say we've done.

Q Now, what I want to do formally is identify all the various documents being read, picking up some caveats about them as we go, and asking whether you adopt them as part of your evidence, because clearly there are hundreds of pages, and we've read them, in some cases many times, and that's prompted the questions, but we don't want to go through them where I will be here all week. So there's the main report, which is dated from July 2024, which is – and this is really for the record – bundle 44, volume 1, document 1 at page 5. Then you produced a response about your calculations, which I wonder if I can look at. It's bundle 44, volume 2, document 3. Now, how did this document come about? Did we ask you a series of questions and then you responded?

A This is our response to the five reports. Is that what we're----

Q No, I think this one might be the response to us asking you what your data was. If you look at, for example, the question 1, paragraph 1, page 80 data mentioned in relation to Queen Elizabeth Hospitals, you're telling us what the data you used were, and then you're referring to various documents which end up in the bundles.

A Ah. This is your questionnaire to us.

Q Yes. And so you wrote this, and we put the documents in the bundles.

A Yes.

Q And then there's the HAD response document itself, which I think

you're responding to, which is volume 5 of bundle 44, and it is at document 2 at page 20, and this is a joint document of 19 July.

A Yes.

Q Yes. And then after that, there is some calculations that I asked you to carry out with Mr Mookerjee, and that is bundle 44, volume 7, page 51, and you've included in that quite a lot of extra coding and results in the documents which follow that, and I appreciate that I asked you the questions and so you have some thoughts about whether the questions are good ones, but is that your document?

A That is my document, yes.

Q And then we asked you two questionnaires. The first one is HAD Questionnaire 1, which is in bundle 44, volume 2, document 1, if we can go to that, and that is page 12 of volume 2. Yes. And so we asked you a series of questions, some of which we directed at just you and some we directed at just Dr Agrawal and you answered those, and you're nodding. Now, there's a person doing a transcript and they can't see nods.

A That is correct.

Q Yes. And then finally, there was the second questionnaire, which came in after we got other people to review your work, and that is in bundle 44, volume 5, document 1 at page 4, and of course that included the request that results in the response document.

A That is correct.

Q Right. Now, clearly there have been some changes, and there's some things I asked you to do which you need to express in view on whether they were the right questions to ask. But in broad terms, are these part of your evidence, or are you happy to adopt them as part of your evidence?

A Yes.

Q Yes. Now, what I want to do first is think about how you were instructed, because if we go to your letter of instruction, which is bundle 44, volume 1, document 5, page 225. No, it is page 239. It's not even that at all, sorry. It's not. It's 244. We're in the wrong volume. It's volume 1. There we are. That's the right place. Right. So, this is a letter which we're provided by the Central Legal Office acting for Greater Glasgow and Clyde Health Board dated 8 February 2024. Is this your letter of instruction?

A This is my letter of instruction.

Q So before you got the letter of instruction-- I'm assuming this just didn't appear in the post randomly. Had you had discussions with Dr Agrawal and Professor Hawkey about the possibility of being instructed?

A I did. I first had a very brief

discussion with Dr Agrawal. We met together with Dr Agrawal, and I don't believe Professor Hawkey was there. We met with [redacted] to discuss what----

Q This is a lawyer from the Central Legal Office.

A That is correct.

Q Okay. So, your letter, do we see at the bottom of the page, the last big paragraph, it says, last sentence:

> "Thank you for confirming that you are able to assist and please liaise with Dr Agrawal and Professor Hawkey as to the specifics of the statistical analysis required."

And earlier in that paragraph, it says on the third line:

> "A copy of the original letter of instruction to Dr Agrawal and Professor Hawkey is appended...."

So did you read the letters to them?

A Yes, I did.

Q Right.

A Well, I read whichever was appended first, because they look the same.

Q Yes, of course, because they're the same. Can we look-- I mean, I don't know which one-- So let's look at the appendix to their letter. So the nearest one is page 237. So the letter describes this as the questions that are to be asked. Now, what I'll do is I'll ask Professor Hawkey and Dr Agrawal about their letters themselves and just focus my questions to you directed at the questions. Now, should we understand these are the questions you were effectively helping to answer?

A Not all of them, but yes. So for example, the first question, "Are hospital water systems sterile?", Professor Hawkey had done a lot of that work already by the time I came on board and they wanted help with the statistics and the data.

Q That was what I was going to ask you. If we look at these questions, are there any of these questions effectively that you didn't contribute to in that-- Although you obviously read the results, but in terms of doing the practical drafting, is there anything I should go and direct my questions elsewhere in their entirety?

A Yes, so the sterile hospital water systems, that can go to Professor Hawkey, I think.

Q The second one?

A The second one I didn't really address either.

Q But then the rest of them possibly. What about 3, do you think that would have been one you would have been involved in, the water testing?

A Yeah, no, I didn't go through any of the documents on water testing or anything like that. I focused predominantly on what the data might tell us.

Q Okay. So effectively, we'd see that within the answers to 4 and definitely 5 and then 6.

A Yes.

Q And definitely 7. Right. The reason I ask that is there is a document referred to in a footnote to question 3. So if you see on the third line of question 3, there's a small suffix number 1 and a footnote at the bottom.

A I do see that.

Q With a web link. Now, that is a document that the inquiry is familiar with, which is in bundle 18, volume 1, document 11 at page 819. Now, did you read this document or until we started instructing you, had you read this document?

A I had not read that document until you instructed me.

Q The reason that I raise this immediately is because it prompts a question. I think What it does is it describes the December 2018 perspective of Health Protection Scotland and what had happened since the start of what is known as the water incident in February/March, and it contains a lot of facts, and it uses the concept of water contamination, it discusses what they think's going on, and it provides all the context. To what extent would you accept that not reading it has an effect on your understanding of what you're looking for in the data?

A I think that's an interesting and fair question. I would argue there's a positive and maybe a negative to not having read it, the first being sort of objective about the data instead of having a biased opinion before I start with the data. That's the positive side. I think the negative side, of course, is when you're trying to interpret what you see you might want to understand what's happening.

Q Right. How should the Inquiry-- Because I think it's fair to say that when we ask you the first questionnaire – I think it's question 10 and 11 – we give you a long list of documents, and then question 13, there's a long list of facts, and these are everything from the DMA Canyon report through that summary into all the epidemiology that we had and various facts about the state of the water ventilation system. Your response as a team is, "We didn't look at that," or, "It was outside our remit." You'd accept that?

A I would accept that.

Q How is the Inquiry supposed to react to the report you produced before you became aware of that information,

the HAD report, when it is produced in ignorance of all that context? What are we supposed to do with your conclusions, considering they were reached absent the context?

A So I guess, again, the positive comment I would make is, you now have a set of data and, you know, a set of questions from completely objective people who have not been involved in any of this. I think that that's actually very helpful sometimes, when people sort of become very narrowly focused on a concept. Does it help to then learn later and interpret? Probably, and I think that's what we have done largely with you, learned about some of these later. You pointed us to certain areas and we can consider our findings in light of that information.

Q So I wondered if you'd accept that you are being asked in these questions to consider in a sense what would be seen in the data if there was widespread water contamination of the hospital?

A That is correct.

Q Yes. Did you know when you wrote the report what the widespread contamination was thought to be?

A So, we had some awareness. You can see in the portion of the report that was largely written by Professor Hawkey that he saw CNR details about what micro-organisms were found in the water. We also had the whole genome sequencing data, and we did look at those.

Q Okay. In terms of the ventilation system, would you accept that you were being asked to consider what would be seen in the data if there was inadequate ventilation?

A Yes.

Q To what extent did you know what was being suggested as being inadequate about the ventilation?

A In terms of adequacy about the ventilation, I think we were-- we felt that we were being asked if the ventilation didn't meet the NHS standards of 10 air changes an hour, and I think there was some discussion about HEPA filters as well.

Q Do you know what the state of the ventilation system is in the various wards of the hospital?

A Well, I'm well aware now.

Q But you didn't when you wrote the report?

A I think we were aware that it wasn't to standard, but we didn't know exactly the details.

Q Because it doesn't say that in Chapter 6. There's no discussion in Chapter 6 of what's wrong with the ventilation. Why is that?

A So Chapter 6 was largely led

by Dr Agrawal. I did the analyses-- Or I supported some of the analyses. I didn't really get heavily involved until you asked us questions later and you wanted rates----

Q Right.

A -- but, yeah, I don't-- I would refer you to Dr Agrawal for that.

Q So just to be clear, before we pass it by, Chapter 8 in the HAD report is a Dr Agrawal product, largely----

A Largely.

Q -- but your work on Aspergillus is, you then got involved once we started calculating rates?

A Yeah.

Q Okay. Now, I'm not sure what the right word is to use for this, so I'm quite keen to see what you think is the appropriate phrase, but would it be fair to say that the two scenarios you were considering were widespread water contamination and inadequate ventilation?

A That's fair to say.

Q Is scenario the right word, or would you get a better word that's more appropriate for epidemiology?

A No, I think-- I think that's fine. I think we can go with that.

Q All right. Now, you've mentioned to me when we've consulted the concept of the counterfactual, and we've evidence about that from a number of people in evidence. Before we go on to look at the rest of the report, I thought it'd be a good idea to see what you have in your mind as counterfactuals that we ought to consider. So, in order to get onto the counterfactual list, would it be fair to say that a scenario has to be not implausible? Is that a right way of putting it or is there a better way of putting it?

A I'm fine with that.

Q Okay, right. So, what are the not implausible counterfactuals, as it were, that you would like us to consider alongside the scenarios of widespread contaminated water system and inadequate ventilation?

A So the first one would be line care, of course. That's the first thing you're looking at when you're looking at bloodstream infections just in general, long before I would have had any about them working on their line care. That would be my first-- my first go-to. There's a mention of nursing shortages: again, you get rushed behaviour and your line care can suffer. There is some suggestion of teams not working together well: again, you get poor handover, you get less safe care. The other thing one might explore – again, I would ask you to explore this with Dr Agrawal – is anti-microbial usage and whether or not that's driving up certain micro-organisms because of----

Q So we recap: there's line care, nursing shortages----

THE CHAIR: Is nursing shortages a separate counterfactual, or is it----

A I would say-- I would say it is. I would say it is, and alongside the nursing shortages – I apologise, Fred, for jumping in – when they move from what I think are probably larger bays to single bed bays, that creates a problem too. You need more nurses to cover single bed bays than sort of a four-bed bay or a----

MR MACKINTOSH: Now, why do you say you need more nurses? What's your professional experience that tells you that?

A So my professional experience is working with the Infection Control Team at Imperial, and this is-- and also with nursing-- with the head of nursing. This is a commonly held belief, and the way they run the NHS, the-- when you have rooms that have single patients, so a single-bedded bay, you need more nurses because patients need to be in the visual field of nurses, to a certain extent need to have the nurse there responding to them, even if they're, you know-- they're still looking after more patients. So, "In a larger bay, you need less nurses," is the NHS' assessment. I am not going to challenge nursing in the NHS----

Q So you would see "nursing shortage" and "single rooms" as related to each other? As in----

A Those are related, yeah.

Q Those are related? Right. Then you have teams not working together.

A Yes.

Q Are we able to work out which teams you have in mind for this counterfactual? Because I mean, there could be some candidates. So, I'll give you some candidates. So you could have Estates people not working together----

A Yeah.

Q -- you could have the higher hospital management not working together, you could have the Infection Control Team not working together, you could have the Treating Clinical Team not working together, and I'm sure there's others I haven't thought of. In that sort of context, are you looking at a melange of all of them, or any particular in mind?

A I would say the entire melange really can have different impacts on patient safety and patient care.

Q Okay and then anti-body resistance, I'll discuss it with Dr Agrawal.

THE CHAIR: Right, just for my notes could I just identify which teams are in the melange from your perspective?

A So-- And this is-- I-- I want to highlight that it isn't clear, it just sort of

feels suggested in the documents that I read, but-- that there can be tension in Infection Control and the nurses and the doctors that are on the ward. There can be tension with Estates and the people who are supposed to be maintaining that estate, and then less so that I see anything about management necessarily, but oftentimes this comes down from management and how management is working with their teams.

MR MACKINTOSH: Now, I'm going to come back to these later and ask about, in a sense, the mechanisms that you have in mind that put them on the "not implausible" list for both those and the scenarios, and then we'll look on your opinions and we'll try and, in a sense, break them down into manageable chunks. What I want to do before that is to look at the data----

A Okay.

Q -- because it figures that if we look at the data first, at least we have something to hold on to if we get lost. I think you've already covered this, but I want just to be clear: in the HAD report itself, should I be focusing my questions on you for Chapter 7?

A Yes.

Q And effectively leaving Chapters 1 to 5 to Professor Hawkey, Chapter 6, Dr Agrawal, and Chapter 8, as it stands, to Dr Agrawal.

A Yes.

Q Right. I'll do that, but when we just look at Chapter 7, do you recollect it has three parts? The first part is about water testing results, and the third part is whole genome sequencing?

A Yes.

Q Within that division, where do I go-- The middle part being the data, do I go to you for the middle part and the other bits to your colleagues, or how do you feel this splits?

A So, I am happy to answer on all parts. I may defer some of the whole genome sequencing to Professor Hawkey.

Q Right. Now there may be a question of time, and I might leave whole genome sequencing to the end and take a slight canter through it and put to you some broad propositions and see how you respond, but leave the hard work, as it were, to the professor. What we'll do is, if we could look at the declaration to the authors-- So that's on page 151 of bundle 1. So you gave a declaration. Now, have you given evidence in court proceedings before?

A I have not.

Q So is this the first time you've come across something like this, or is there something similar in academia?

A There are-- Yeah, there are similar, but not quite this legal of

declarations in some of the other consultancy that I've done, which isn't the same thing.

Q I've got two questions, one of which is broad and one of which is precise. The broad question is, could you summarise in your own words what you feel the principal duties of an independent expert are?

A Yeah, and I think I summarise these for the CLO's legal team. I wanted to meet with them before I agreed to do this because I wanted them to be clear in-- my ethical point is that, as an objective expert, I'm giving an objective opinion, and if the data show something they don't want to see, that's still what I'm going to report. So there is no-- So I think an expert really uses their expertise and the knowledge that they have, and tries to stay completely objective. We aren't working for someone, in that respect.

Q The precise question relates to Question 6.

A Okay.

Q Are you confident that you and your colleagues have shown all the sources of all the information you've relied upon in HAD 1?

A So I would have said yes, but I'm going to say no now because of what your team had requested, which were data sets and other things like that, which I was happy to provide.

Q Because in essence, for reasons-- I think you're going to explain why you did it, but for reasons you have explained in the notes, your sections of the HAD report don't contain tables of occupied bed days for the children, and they don't contain coding and calculations. I mean, you'd accept that?

A I would accept that.

Q Why didn't they contain that?

A So they didn't contain them because normally when we show data, we want to describe it, and having just data sets are not seen as helpful when you're describing patterns of data. However-- And we believe that, sort of, the data were shared, and so in hindsight I would say that was a mistake on my part. I should have ensured that the data sets were all shared and ensured that with the CLO, that they had indeed shared the data with you and you had the data.

Q So what I want to do is move to Chapter 7. So if we go to the beginning of Chapter 7in the report and we look on page 62. Now, the questions I had about 7.1 related to what guidance exists for the Scottish Health Service in management of water systems, and whether this approach is consistent with that guidance. Is that something you feel you can help me with?

A I'm going to defer you. This--

7.1 was written by Professor Hawkey.

Q Well, I think it's probably better we do that with him.

A Yeah.

Q So let's move over the page into 7.2. Now, the question, and this is a stylistic thing and you can tell me whether I'm being unfair, whoever designed the headings for your document seems to have a habit of putting questions-- things that look like questions in the headings but they don't appear in the text below. Is that something that you've noticed?

A I found it quite difficult, to be honest, to get the analyses. So I did a series of analyses based on discussions that we had and I felt it a little bit difficult to get the sections and write up to my analyses under the headings, but they weren't my choice of headings.

Q Right, because the question I wanted to ask about 7.2 is-- Now, firstly, you must correct me if I've misunderstood the process, the logic behind this. Are you testing a hypothesis here, effectively, in this section, the hypothesis being if there was widespread contamination, we would see a signal in the data? Is that what the question-- you're asking yourself?

A That was not necessarily the question. I would nuance that question a little bit and I would say----

Q Okay.

A -- if we look at the data and we look at the patterns of infection, could that be explained by contaminated water? So I would reverse that a little bit.

Q Okay. In order to answer that question, do you need to know when the water is said to contaminated?

A Yes, and it was our understanding that it was from the beginning when they moved over.

Q Because that's not what the HPS report you haven't read says. I would summarise it, and I'd rather to save time, as at the beginning there wasn't a problem with the water. It's only until '16/'17 that people become aware of it, and if you haven't read the HPS report, how can you look for the impact of that contamination if you don't know when it was?

A I would agree with that.

Q Right, okay. Obviously, you've looked at both adults and children in this section and the first thing that happens in this report, I suppose, is the sub-creation the environmentally relevant and non-environmentally relevant groups of organisms. Is that your work or somebody else's?

A No, that's mine.

Q That's yours, right. Why didn't you follow the classifications that had already been used?

A So when we were deciding

what was environmentally relevant, Professor Hawkey, who is an expert in this area and has worked his entire life as a researcher and clinician in microbiology, predominantly with bacteria, went through each one. I produced the entire list of species for him. He went through them one by one and made calls on every single one of them of whether or not that is something that can persist in the environment and then potentially infect a patient.

Q I appreciate that and I'll ask him about that in detail, but the sort of practical question is this, it's this report was produced in the summer before the third hearing of this Public Inquiry. Some years after the HPS team had attempted a number of different classifications of organisms, and perhaps the most common split they picked, for better or worse, is in bundle 7, document 6, and I'm not going to do the page from memory because that's going to be dangerous, is page 219.

THE CHAIR: With apologies, Mr Mackintosh, for my notes?

MR MACKINTOSH: Bundle 7, document 6.

THE CHAIR: No, I've got that. It was really a previous point with Dr Drumright. Dr Drumright gave the filter which Professor Hawkey applied to your list, and it's just-- could I take that at dictation speed again?

A Yes. So I produced every single species there were in the microbiology data set that we received, and Professor Hawkey went through each species and, based on his knowledge and also of the literature, he has a very strong command of it as well as his knowledge of the microorganisms, decided which ones could persist in the environment. I think because the question was water could also -- could persist in the water and then would subsequently have the potential to cause an infection in a patient.

MR MACKINTOSH: So I appreciate that he's done that from first principles----

A Yes.

Q -- and I'll ask him about why he did it from his point of view, but this is just more about the practicals of epidemiology----

A Yes.

Q -- because one of the problems this Inquiry has to do is it's got, I've count now, but somewhere about seven different attempts at this and if we look at this, this page is from the October 2019 HPS review. For better or for worse, they spent some time reviewing the rates of infections for gram-positive bacteria, gram-negative bacteria, and what they self-defined as an

environmental bacteria group and such group where they added some enteric organisms.

Now, they had excluded fungi for a reason they've given. Now, I'm not saying they're right by any means, and that's a question more for Professor Hawkey, but might it have given more utility and made it easier for this report to be used if it had connected to somebody else's report as well, because one of the problems we have is comparing apples with pears?

A And I appreciate that that comparison is difficult but I-- again, I'll refer to Professor Hawkey, but I would say, you know, if these are not great classifications or they're not an objective classification, then why not look at a new classification, right?

Q Okay. Right, we'll take off the screen. What I want to do is to understand how you-- what your sources of data were.

A Yeah.

Q So, as I think you know from the NSS response and other people, there are lots of questions, but you're not the first person to have been asked lots of questions about your numerator and denominator. It is a theme of this Inquiry. So, let's think about your numerator data, your BSI spreadsheet.

A Yeah.

Q You had two spreadsheets, one for adults, one for children?

A Yes, I did.

Q And you had, I think, 25,000 of the children, and I can't remember how many rows there are in the adult, but they're big spreadsheets.

A Yes, they are.

Q Right. Now, the Inquiry's experts, only for the new Royal Hospital for Children, they received all the test results for the whole hospital with 215,000-- 14,000 rows, and they then selected from that what they wanted. You seem to have had a pre-selected sample. How did it come to be that you, in effect, didn't get to choose which infections went on your list?

A So when I joined, which was later than Professor Hawkey and Dr Agrawal, there were data sets----

Q Right.

A -- sitting in file that was shared and I was instructed that those were the data sets I was to use, and the microbiology data set were I was told were from the haematology oncology patients, both adult and pediatric.

Q Because one of the challenges that's made to your methodology, which it may be that the wrong person to ask about this, is that what HPS claimed they did, and what Dr Kennedy claimed he did, and what Mr Mookerjee and Dr Mumford

and Ms Dempster claimed they did, and there's a big argument about whether they succeeded, but what they all claimed to have tried to do is to identify patients who were in Ward 2A, the Schiehallion Unit, and its successor units and try and understand what happened to infections in those physical geographical locations. Is that what you understand they were trying to do?

A That is what I understand, yes.

Q Yes, but you're not doing that, are you?

A No.

Q No, because it's been suggested by Ms Cairns yesterday that, effectively, what you're doing is asking what the rate of infections were for people who have a paediatric haematology-oncology consultant in the children's.

A Again, I can't confirm if it's based on the consultant or based on the patient.

Q Because the data you seem to have seems to be limited to people who are flagged in the system as haematology or oncology. You're nodding again for the person doing the transcript.

A Yes, I would say that that is my understanding.

Q Yes. So how do you react to the criticism that that creates a problem because you're no longer looking directly at the physical space, which is the subject, in the case of the paediatrics, in the interest because people are saying, "We are concerned that the widespread water infection-- water contamination is causing a problem in the Schiehallion Unit."

A Yeah.

Q And you're not asking that question, are you?

A No.

Q No.

A We're looking at water contamination or widespread BSI across that patient population and that's what we're looking at patterns of. So we aren't asking the question of the Schiehallion Unit, and----

Q So my next question-- sorry, carry on.

A -- sorry, and what I would say is, again, this gives you more information, whether or not, you know, it's for the to decide if it's helpful, but it's not so narrow to the one unit.

Q Right, so----

THE CHAIR: Just so I'm following, when you say, "Not so narrow," you've been provided with a list of cases described as haemato-oncology cases. You accept that's accurate.

A That is accurate.

THE CHAIR: But you also accept that it is not necessarily a population

relating at which relates directly to treatment in a specific location?

A That is correct. That being said, if we believe that the patients are in the location they're in, and I think you're going to ask me about that too, on the microbiology report, you'll notice that most of those infections are coming from patients in those units. Same thing when we go over to the adults. It's BMT, right? So the most fragile patients are carrying the greatest burden of BSI, which is not a surprise.

MR MACKINTOSH: Because we have a series of patients who have given evidence, and their families have given evidence, of how paediatric haemato-oncology patients end up on adult wards or other children wards and get infections and, again, you're nodding. Now, the question that I want to ask you is this.

A Yes.

Q Is there a risk by your-- the approach that you've been given that you include those cases, because they are haemato-oncology patients, but in a different context, because they might be exposed to a different ventilation system or a different water system.

A That is correct, but if they were on that ward to begin with, the Schiehallion ward to begin with, they were also exposed to that ward, and as we get into how someone can get a BSI from exposure to water, for example, that becomes important.

Q Okay, and then, well, we move to the denominators. There's two questions, one is conceptual and one is numerical. The conceptual is you received the denominator data set. Had that arrived before you did as well?

A It did.

Q Right, and you were told that that matched-- well, what were you told about it?

A I was told that was-- those were the denominators, right? I asked for bed days per month, per year, and I was told that's what they were for the settings that I was looking at and for that patient population, and because it was given to me by the NHS, I took that as the truth.

Q Because one of the problems that other authors have got into is that they have attempted to match the occupied bed days or the admission days to Ward 2A and have found it harder than they thought it would turn out to be and have been criticised for that. So what I'm suggesting is that if you are including patients who are elsewhere in hospital when they have their blood test, but you're only including the bed days in the Schiehallion Unit, and that's what you think you're doing, then again, is it that you don't?

A That's not what we're doing.

Q All right, okay.

A So those are not bed days for the Schiehallion Unit, those are bed days for, as I understand it, haemato-oncology patients in each of those hospitals or sectors as the case may be for adults.

Q The question flows from that is that if we look at what you're measuring, you're trying to measure the rate of infection amongst haemato-oncology patients irrespective of where they are.

A Yes.

Q Although most of them are in the Schiehallion Unit, you think. And you do that by dividing by the bed days of all the haemato-oncology patients. Is that effectively what you're saying you're doing?

A Yes.

Q Yes. If you have one haemato-oncology patient – and this is an extreme example, but I think it might've happened – who is sitting on a ward in the children's hospital where there are no other haemato-oncology patients and you want to measure their risk, do you not need to take account of the activity in the whole ward they're in?

A So, normally, yes, that's the way that would be done. So, for example, if they're out in, I don't know, a gastroenterology unit - let's make it very extreme - you would want to know normally in surveillance what ward they are in and their exposure there, right? So that's your denominators in that ward.

But, you know, I understand the limitations of the NHS where patients are put in wards that aren't part of their healthcare area.

Q So, the question that arises is, to what extent is what we've just discussed disclosed in the HAD Report, chapter 7? The fact that you are looking not at the Schiehallion Unit, but at the patients spread around, wherever they happen to be.

A I believe we said that it was the haemato-oncology patients, not the Schiehallion patients, and we have tables and documents listing clearly other wards besides Schiehallion.

Q You've seen the final document from Ms Cairns in which our request pulls out the numbers from the HPS report that we were just looking at, that show the occupied bed days for the Schiehallion Unit and its predecessors. She points out that your numbers bear no relationship to that at all. I get the impression that your response is-- Well, what is your response to that?

A Well, I have a couple of responses to that. The first one is it's quite difficult for both the Inquiry but for anyone trying to help in this space to be able to-- or monitor healthcare infections, to be able to do so, if we can't even get a

standardised denominator, right? So I would hope, at a minimum, we're talking about standardised datasets, really important to the NHS, really important to surveillance, really important to patient safety. We could just get a standardised denominator.

That being said, and the questions you're currently asking, I wonder if we are comparing the same denominators or not. So I wonder if-- I did see Ms Cairns' document. I was pretty surprised that it was that different, given that-- where she says her data comes from, at Health Protection Scotland. It should all be coming from the NHS. I received my data from the NHS. I just wonder if those are different-- you know, they're pulling from different wards. I wonder if it's from different wards. I don't know.

Q Because the question that seems to arise in this is, I absolutely get that you're of the view that, however this is different, it's still interesting and useful. We should look at it. I'll park that to one side and just look at the negatives for a moment.

You were supplied with occupied bed day data by Greater Glasgow and Clyde Health Board.

A That is correct.

Q They supply, according to Ms Cairns, the Public Health Scotland dataset that Dr Kennedy and HPS used, and yet they didn't supply that to you.

A That appears to be correct.

Q Yes. One of the factors that happens during the story of these events is that the Schiehallion Unit, Ward 2A and Ward 2B, moves on 26 September 2018 to Ward 6A, unless you require a BMT room, to Ward 4B. That is a geographical location change and one which we're really interested to see whether we can see any consequences of it in the data.

A Yes.

Q And yet you're not actually modeling or understanding the infection rate in Ward 2A or Ward 2B or Ward 6A in your work. You're monitoring a slightly more diffuse community of patients. Is that a fair criticism?

A Yes, I would say it's a broader community of patients.

Q Does that have any effect on the value that we can ascribe to your number set, your dataset?

A So, what I might argue is that as far as I could see – and I'm no expert on all the evidence that was provided, but of course your team helpfully pointed me to a lot of different documents to look at – everyone has been focused very narrowly and, as you say, because they think this is the unit that has the predominant problem, this is where they saw the water issue. Looking at the broader patient population is very helpful for context as

well, I would say, because if you just look at something narrow and you have no comparisons, you don't actually know that you have an increase or a decrease or a problem. All data could be increasing and decreasing at that same rate.

Q What I want to put to you is that you didn't design this piece of work with that in mind because you didn't know at the time you designed it that the interest was on 2A for----

A That is correct.

Q So, you're effectively retrofitting a positive onto the design that you did build.

A That is correct.

Q Yes. This report has arrived with some fanfare and yet you didn't know the temporal extent of the water contamination and you didn't know how your patients you're looking at related to the physical geography.

A That is correct.

Q Does that not raise some questions about whether you designed the piece of work at the beginning very well?

A So, that's an interesting question. I think these were the data we were asked to work with and as you well know when you're consulting, you're asked to do a certain project and you agree to carry out that project. So these were the data we were asked to work with. I would suggest you ask the CLO, why this dataset.

Q Well, we can. Whether client confidentiality will prevent us being told is a different question. I need to ask you the reason that I asked all that, which is that there were two cases which NHS Greater Glasgow confirm were linked to the water by means of whole genome sequencing. One of them is a Cupriavidus case from 2016, and the other is a 2018 Mycobacterium chelonae case. Now, what I'm intrigued to understand is why those two cases don't make it into your exercise at all.

A We do have Cupriavidus.

Q But not the 2016 case.

A Oh, really?

Q And Mycobacterium chelonae isn't in your dataset.

A Isn't in there.

Q Now, you're not the only person not to have had Mycobacterium chelonae in your dataset. The CNR originally didn't at one point. But why is it that in being asked to do this exercise and receiving a dataset, the two cases that the Health Board accept have a link aren't in your dataset?

A That is a question for them.

Q Now before we go back to the children as it were, let us look at the adult cases and try and understand something about the case design that relates to

them. We'll go to page 69 of bundle 1. So this is 7.2.3, "Bloodstream infection to adult patients," and this page has, at the bottom, a sort of summary.

Before we discuss this in any great detail, you've looked at North, South Sector and Adult BMT.

A That is correct.

Q Now, did you decide that division or that predates you?

A Predates me.

Q What do you understand to be represented by those three designations?

A So, my understanding from what I was told is those are the adult haematology-oncology patients in those sectors.

Q Right.

A And that South moved to QEUH, as well as BMT, and those dates are in there and I can recall those dates for you if you want, and that North Sector remains in its location outside of QEUH.

Q What's your extent of your knowledge from-- Well, I mean, you probably know more now, but when you wrote the report, to what extent did you know the extent to which the Adult BMT patients were exposed to unfiltered water in the hospital?

A That, I was not aware of. We were told-- So, if you note in the way I set up the data, when I split it between QEUH and other hospitals representing those sectors, I do account for that temporary move, which I now understand BMT moved in 2015. They were supposed to go over with South and then they moved back again. I did not know why that move occurred.

Q Right.

A And that is now my understanding, having seen the water.

Q Yes, so what I want to put to you is that whilst the North Sector stayed in the North Sector and, as far as we're told, there was no suggestions of anything wrong with their water or ventilation, it seems to be uncontroversial that the South Sector were in Ward 24 in Southern General and then they move into Ward 4C, and they are exposed to the water until the filters go on. Do you know when-- Well, you know now, but did you know when the filters went on?

A We did. This is something I did not know anything about.

Q And the BMT patients, as you say, they're in the hospital for five weeks in the summer of '15, then they go away and they come back in I think June '18, at which point there are filters.

A That's correct. Yes, that's my understanding.

Q So, does it cause any difficulty to the structure of the adult part of this chapter, that the Adult BMT patients, apart from that five-week period, don't

appear to have been exposed to unfiltered water at all?

A So, in hindsight, had I been aware of that, I would have modeled the different sectors separately and I do a little bit of that in the----

Q Well, we might come to that in a moment. Now, Ms Cairns raised an issue with you, which I think has been resolved, and I want to go to her document, simply because I think it's quicker so you can see what I'm talking about. So if we go to bundle 44, volume 2, document 45, at page 693. Definitely not that one, 44. The good thing about this week is it's always bundle 44, unless it isn't. So that's volume 2, page 693, please. Yes, it's paragraph 3.2.5.1.

So there's a discussion here from Ms Cairns about consultant sectors, and you respond to it in your response document. What was the point you thought she was making here?

A So-- Yeah, so she had-- So, I think the point she was making here is that, and this a difficulty with working with somebody else's data without speaking with them, right, is that I had misassigned some of the sectors to QEUH. Now, there was a-- there is a column in that data that's QEUH, "yes" or "no," that's the assignment and that's not the column she was looking at, is my understanding.

Q So, you've misassigned something, she thinks, but actually in the one you actually use, you hadn't misassigned.

A That is correct.

Q Is that effectively the resolution?

A Yeah, I double-checked that.

Q Okay, right. Take that off the screen. What we'll do is we'll go and look at the report itself. That's 44, volume 1, and we'll go to page 73. So I want to just look at a few charts and check I understand what I'm looking at and then ask you some questions.

So, if you go over the page to Figure 5, without discussing what's in it, just in terms of the colours----

A Yes.

Q -- who are included in the red line?

A Okay, so the red line, and I apologise for this not being clear-- So I have the data given to me from North, South, and BMT. So in that red line, anyone from those sectors in the dataset that I'm provided who are not currently at that moment in time assigned to QEUH. So for North, that will be for the entire period; for the South sector, that will be until-- I believe it's May or June 2015; and for BMT, I take that 2015 period, the June and July, they're over in QEUH during that period. Otherwise, they're included in the red until June 2018.

Q So it is a geographical split?

A Yes.

Q And so adult BMT move around quite a lot and South move around once?

A Yes.

Q Right. So if we go back to the previous page, the bottom of page 73, you discuss:

> "Incidence of environmentally relevant organisms by QEUH vs other GGC hospitals (Figure 5) demonstrates similar incidence by month per year of environmentally relevant organisms, with the exception of both QEUH and other hospitals in GGC having period spikes in incidence. Of note, from mid-2016 to mid-2018, there appear to be a series of spikes in incidence at QEUH, however by mid-2018, incidence rates in both QEUH and the other hospitals are similar, if not slightly higher at hospitals. "

And then you say:

> "When we adjust for unrealistically small denominators, predominantly from the South Sector, which moved to QEUH ... we see a dampening effect in the spikes, with a single possible true spike in incidence at QEUH in early 2018 followed by significantly lower incidence for the remainder of the time period. This suggests that there may have been an overall brief spike in incidence compared to background at QEUH at this time, with constant and relatively low incidence following, and even a possible reduction in incidence from 2022 onwards of bacteraemias attributable to microorganisms that can persist in the environment."

And we can actually see that the changed one is figure 6, which is at page 75. So I have two questions about this exercise. The first relates to the unrealistically small denominators. Now, what did you do when you thought there were some unrealistically small denominators?

A So the first question was-- I posed that to the CLO that we were working with, because I said, "Are you sure these are right?" Unless they've closed a ward, I wouldn't have expected to see that low of a denominator in bed days. Then it was sort of like, "Well, that's the data you have," so I said, "Well, I guess what I could do is look at what it

looks like every other year that's not that year in that same month and sort of calculate an average and take a look at that and see what it looks like."

Q So one observation that was made is that in some months there are possibly unrealistically big denominators. I mean, if we go on to page 72, just to pick one, 2014, April, page 72. That's the pop-up there, fifth row down. In South sector, there's 447 bed days, which seems quite high. I mean, I don't know. And if you look down into 2022, there's a similar-- there's a 332 in what I take to be May. Did you do anything about unrealistically big denominators?

A I did not, and I think that was a fair comment by Ms Cairns.

Q Now, I appreciate you might have spoken to CLO and they told you you had the data you have. You didn't feel it was appropriate to ask them to contact the consultants who run that ward who could have told you what was going on?

A I was not aware that that was an option.

Q I mean, we have statements from those consultants which were taken while you were doing your work, but we weren't asked to do that. So would it have helped to know some context from those consultants?

A It would have helped tremendously. I think that, you know, had I designed the study, in hindsight, I would have asked for all the data, I would have been aware that other people had modelled this and asked for their data as well so I could look at data comparisons, and if we were able to talk to the consultants or the nurses that were working at the time, that would have been very helpful in understanding----

Q Because one of, I suppose, the problems with this exercise is-- If we sort of flip inconveniently – and we're not going to do this; we'll just do it in our minds – between page 74 and page 75, if we go to page 74, we have, as you described, between '16 and '18 a series of spikes in the blue; and we go to page 75, and we have less spikes between '16 and '18. Now, what's your view about what we should look at and ask questions about in these two charts?

A So, this was an exercise to just caution that there is what we call noise or wobble in the data. I modelled everything on the data that I received, so I did not use the corrected denominators for most of the data that I modelled.

Q Because there's nowhere in this section any discussion of whether any of these spikes are significant.

A Yes. This is merely a plot. This has not-- So later on, as you're aware, because you requested this, this

does not have models like a GAM model or anything like that looking at whether or not these spikes are significant.

Q But because this is really the only bit of the adult data that posits this question, I'll ask it now. Given that we know from the water testing results, albeit not for this Ward 4C, for other wards-- Well, we actually know for large parts of the hospital there were high positivity. Now, there's a debate about whether the positivities are accurate, and a witness has explained that they should be downgraded for reasons to do with the thresholds, but an argument can be presented that in February, March of 2018, there was high positivity in water testing results. Given that that's a fact of-- well, it's not quite a fact, it's a series of bits of evidence we know, should we not be looking at that early 2018 spike and raising our eyebrows quizzically and wondering what's going on? And indeed followed by the drop-off that you notice afterwards?

A Absolutely. I would say that, again, if I was designing this study, I wouldn't just look at water. I would look at everything that might be going on at that time period that could attribute this spike.

Q Right. So there might be other counterfactuals here?

A Lots of them, yeah, that we've

listed before.

Q Yes. But suppose, in order to give advance notice of the questions I'm going ask at the end about the paediatric data, we listed those counterfactuals.

A Yes.

Q So we've got water, we've got-- We're going to ignore ventilation for the purposes of these-- Is that the right approach?

A Yeah. For bloodstream infection, I would ignore ventilation.

Q So we've got ventilation, we've got CLABSI, we've got nursing/rooms, we've got team dynamics, antimicrobial resistance. Now, do you get CLABSI spikes that just go up and go down again? Is that how CLABSI works, or is it more just what we've seen in the dataset that Ms Rogers talked about where it's a big thing?

A Yeah, I would expect it to be more of what we see in the paediatric data.

Q It's spread over time.

A Spread over time. They discover it; they address it; if they address it well, you see it go down.

Q But it's not something that happens briefly.

A No.

Q No. So can we put CLABSI out of our minds for this spike?

A Yeah, I guess so.

Q Yes. Team dysfunctionality?

A Depends on what's going on and how long it lasts.

Q I mean, it's very unfair that I know the names of the consultants I have statements from and they haven't been asked, but I put it to you that the sort of team dysfunctionality that you probably get a flavour of from the material and we've heard a lot of evidence about and his Lordship's not made his mind up about yet happens over a long period of time.

A And that's normally what we see, yeah.

Q Yes, and does that cause just a spike?

A No. Well, I would expect it not to, depending on what it is, but of course, it depends on the teams, doesn't it now? If it's the estates team refusing to do something that the clinical team wants and then somebody gives in, that's another potential----

Q And then antimicrobial resistance. Maybe not your thing. Does it have any spike weight?

A No, because antibiotics go up over a certain amount of time is my understanding and I had a look at Seán MacBride's data.

Q Well, can I ask you about Seán MacBride's data now?

A Yeah.

Q But I'll do that after I've finished this section.

A Yes. So it just suggests there's a pattern over a long period of time is what I wanted to mention.

Q So what I'm wondering here is-- I know it's a spike. I'm instantly nervous about spikes. Would that be the right approach, to be worried about a spike as a signal?

A So unfortunately, with this retrospective data, I would say normally I'd be worried. I'm a little less worried because I'm more concerned about the quality of the data rather than anything else here.

Q Right. But if we take the spike in this reduction you've observed, what was the----

THE CHAIR: Just so that I'm following, when you are being asked about being worried about a spike, it occurs to me that the worry might be about a number of things. Are we talking about the worry being that something has happened which is demonstrated by the spike? Yes.

MR MACKINTOSH: Can we go to page 74?

A Yeah. I believe it was my understanding, just to clarify, that that is what I was being asked.

Q Yes. So, if we go to page 74, what you've said here is:

> "...we see a dampening effect in the spikes, with a single possible true spike in incidence at QEUH in early 2018 followed by a significantly low incidence for the remainder of the time period."

And I'm just wondering whether I can put this to you: that whilst one should be careful because it is a spike in some retrospective data, to what extent is this not inconsistent with a suggestion that the water in the hospital was problematic at that time?

A So, this is where we start getting into the biology of how someone gets a bloodstream infection from a microorganism that might reside in the environment, and that's also not super likely to cause a spike, unless of course there was an outbreak from a common source. So typically, you see a solution that they're injecting in patients becomes infected or something like that, a tool that they're using becomes contaminated, and then you get this sort of single source. You still wouldn't see it go up in one month and come straight back down----

Q Unless you put a filter on the taps at the end of the month.

Well, here-- But this is the issue that I have with that. When you put the filter on the tap, and let's say we're getting part of the contamination by bad line handling and they're contaminating the line from the water that they've touched or the environment they've touched, that's only going to be part of it. The predominant way patients get these types of bloodstream infections is of course by colonisation in their gut.

Yes.

A Right? And that's going to take a long time to dissipate because the colonisation in your gut doesn't go away immediately when you change the tap.

Q Well, indeed, and so this is one of the things that's been put to a few of the experts about enteric organisms, and what you would see, and I think it will get put to Professor Hawkey, assuming I remember to do so, that if the problem in a particular patient is that they have a enteric organism that has got into their bloodstream from their gut, the frequency of those in the unit won't really change over time because it's to do with the sort of treatments they're on, the type of their neutropenia-- Professor Stevens talked about it a great length, but the nature of their vulnerability, and that amongst the cohort of the patients won't really change. So you won't see temporarily restricted increases in potentially enteric bacteria caused by gut translocation because it's a problem that's always there, always happens, always has to be managed. And so, again, conscious it's what you call a true spike, is gut translocation

going to cause a true spike?

A No.

Q Right. But I think I've probably dealt with this one enough. I'd like to briefly about this paper that you mentioned from this pharmacist. Now, what I want to do to see it first is to look at a list of documents, because I'm slightly intrigued by it. If we go to bundle 44, volume 1-- I'm going to have to do this on my computer.

(After a pause) Document 2, page 224. So, what the Inquiry team did at the point when your instructions were novated over to the Inquiry is we received what we rather prosaically referred to as the "data dump", which consisted of all the documents which the CLO for GGC told us were in the working space that the three of you had access to. We see a range of things, and if you see, for example, at 7, we see the first of the Aspergillus diagnostic imaging that Dr Agrawal looked at. If we go down to-- There's a lot of those. If we go to 228, we start seeing your data.

A Mm-hmm.

Q I think it's 48, 49 will be part of your data set. Over the page, there's more of your data.

A Yes.

Q Then we have the overview report for the CNR at 53 (sic)-- No, sorry, back to page 229, number 53 on that list.

CNR overview report, what you've read.

A Yes.

Q We have the Oversight Board timeline. Did you read that?

A No. So the CNR was read by Professor Hawkey and dealt with by Professor Hawkey. I did not read those.

Q Did you read the GGC Oversight Board timeline? Document 55?

A No.

Q The Oversight Board final report at 56?

A No, I did not read those documents. I had a tremendous amount of work to do with the statistics and cleaning up the data.

Q Because if we go on to page 230 and we look at document 86, you said to me a moment ago that you felt you'd like to have seen other people's statistics. There's the October 2019 HPS review. It was in the data file. Did you read it?

A No, I did not.

Q Right. On that note, we go back to page 224 to a document called-- to Row 2, "Public Inquiry Analysis of Microbial Prescribing". Now, I'm not going to take you to it, and the reason I'm not going to take you to it is because we haven't called its author as a witness. I don't think you read it before these bundles were created.

A I did not.

Q Yes. I also know that Professor Gibson, who ran the unit, didn't see it, and it wasn't shown to Professor Stevens when he gave evidence last year on anti-microbial prescription, or Dr Mumford when she gave evidence last year, anti-microbial-- and you didn't use it for your report.

A No, I did not.

Q So, if you didn't read it, it might be interesting, but I'd have to re-run about 15 days of evidence, and it's bad enough as it is already, so I don't want to go to it.

A Okay.

Q But what I want to suggest to you is this: what do we do with the fact that you and your colleagues had this folder, which you must have been going in and out of to take data, and in it were documents that are, superficially anyway, relevant?

A So my-- So I believe-- And you'll have to confirm this with Professor Hawkey and Dr Agrawal. I believe they read a lot more of this than I did. I was given permission to download the data on the-- and hold it securely, the data on the infections, and-- because I had to run my statistical packages and they couldn't upload them to the site they had, and so I predominantly worked with that.

Q Did you look at the Public Inquiry data in Row 1?

A I did not.

Q Because that would have given you 214,000 rows, every single bloodstream infection in whole of the Queen Elizabeth in 2015 to 2023 or possibly 2022, depending on the version.

A To be honest with you, I would have to go back to their data file that they gave me. I'm not sure 100 per cent of this was in that file or not. I don't know, but I didn't go back, of course, when we moved over to work with you.

Q What I want to do now is just to look at a couple more adult haematology issues, just because I think we should probably just pick them up, which is back to bundle 44, volume 1, page 92. So this I think you've already dealt with, but it's the bone marrow treatment patients.

A Yeah.

Q So, these are the patients who you have five weeks in the new hospital, and they go back, and they return once the filters are on. So it says here,

> "We are interested in examining single patient populations as services transitioned from another hospital to [Queen Elizabeth] to determine if there was an increase in the number of cases of

environmentally relevant bacteraemias, which could be indicative of environmental sources of infection."

What are the environmental sources of infection that you were considering when analysing this group?

A So again, this was just the water.

Q If we look on the next page, we have two charts, the top being environmentally relevant bacteria and the bottom being, if we just zoom down to see it, the ones that may cluster. What I'm really asking is, what's the value of this if they were only exposed to five weeks of unfiltered water in 2015?

A Yeah-- And so that was not the information that was provided to us. So it was our understanding that we were just looking at the entire QEUH environment and that the water overall in that environment was contaminated. So, you know-- I mean, what-- what this happily shows is that your filtered water in QEUH looks like it's doing a pretty good job.

Q Because the other thing that we have-- If we look at the top chart for the moment, we have evidence from the Health Board which, to be fair, in submissions to his Lordship, I've said we should accept, and we have submissions from the Health Boards agreeing with me, and we have submissions from other people disagreeing with me, that by '21, '22, '23, the water management was getting much better, they were spending a lot of resources on testing, and potentially we see that in this data.

A Yeah.

Q Would you accept that, the right-hand couple of years?

A I would definitely accept that.

Q So does it does it cause a problem that you, when you wrote this report, didn't know – I mean, it's an arguably vague point in time, but – there's a point in time when the water gets better?

A It would have been very helpful to know that BMT is in a setting where the water is better. So, in other words, they have a move from their other hospitals to-- or from North, I guess, to QEUH, and we don't suspect there's a problem with water all together, except for that six weeks. It would have been nice to compare them to South, where we do know they were exposed to-- potentially to water----

Q Yes. I suppose all we have for South is that previous chart----

A Yeah, yeah.

Q -- that we looked at with the spike, and we've had that conversation. So----

A Yeah, so in hindsight, I would

have modelled that a little bit differently.

Q So you might have had a chart like this just for South?

A Yes, and-- and been able to compare both of them. Right? Because you have North, who never moves----

Q Yes.

A -- and you have-- So all three are a nice comparison. You have North who never moves, you have BMT who moves to presumably cleaner water, and South that's exposed to water that-- that might be more contaminated.

Q Well, just for completeness, there's a last question before the coffee break. If we go back to Figure 6, so that's on page 75, at the point-- We'll actually use the two spikes as our reference points. The point of the right-hand spike, "Other Hospitals" in 2022, that is North and just North.

A That is just North in 2022.

Q Yes, and if we look at the remaining-- the true spike, as you call it, in late '17, or round about the turn of the year----

A Yeah.

Q -- that is only going to be South, because BMT is still under North at that point.

A That is correct.

Q That's really helpful. What I'm posing to do, my Lord, is this might be a good point to break for a coffee, because the next topic is to return to the paediatric PSI.

THE CHAIR: We'll do that. Dr Drumright, I hope we can offer you coffee, and can I ask you to be back for ten to twelve, please?

A Absolutely.

(<u>Short break</u>)

THE CHAIR: Mr Mackintosh?

MR MACKINTOSH: Thank you, my Lord. Dr Drumright, before we go and look at the actual charts and calculations you produced, I want to just return to bed days for one topic.

A Yeah.

Q So we can go to the HAD response document, chapter 2. So that's bundle 44, volume 5, page 32. Now the top paragraph, 2B.3, I think is you dealing with Ms Cairns' comment that we dealt with before, is that right?

A That is correct.

Q Yeah, and then I want to look at-- and 2B.4 is also Ms Cairns' comment.

A That is correct.

Q Now, at the bottom of 2B.3, do you see how:

> "Additionally, the NSS Report suggested that some additions of bed days may be missing with North and BMT sectors in 2013 and 2014 (see

3.2.5.7), we have double checked this and corrected an error."

Now, what I wondered about that was does that mean that the error is corrected in the charts in the response document, having not been corrected in the charts in the HAD report?

A Can I have a moment to read the full paragraph so I can orient myself?

Q Yes, of course.

A Thank you. (Pause for reading) I do recall this now. I would rely on the data that we have provided in the response document.

Q The response document is the one that should be more accurate?

A Yes.

Q Right, and that was what you-- well, in fact, you and I have called in the conversation the "long charts" that we're going to come to?

A Yes.

Q Right. Okay, and then 2B.4 is you dealing with Ms Cairns, we've dealt with that. Let's look at 2B.5. So what was the date range of the BSI data you were given for Yorkhill?

A So I was given data from 2005, January, through to-- I believe it's December 2022 for the paediatric.

Q And what was the bed data you were given for Yorkhill?

A I was given January 2008 through December 2022.

Q And so how are you able to calculate incident rates in 2005, 2006, 2007?

A So I imputed the dates-- sorry, the bed days for the first three years based on all of Yorkhill's data.

Q So, effectively, it's an average?

A Yes.

Q Does that inherently involve an assumption that there was no change?

A It does.

Q Do you have any evidence there was no change?

A No.

Q How do you deal with the-- I think it is a fact that all the doctors and nurses who worked at Yorkhill effectively can't remember that long ago, and we don't have minutes for meetings back that far, and in fact the Health Board, I think, have been through at least one reform since-- after that date. So why is it reasonable to make such an assumption for such a long ago period?

A Can we-- and I don't have it to hand, can we jump to the data provided on the bed days for the paediatric patients?

Q Yes. I think if you give me a moment to find it, we can do that.

A Thank you, because I believe it was fairly consistent for the remaining-- remainder of the time.

Q So if we go to bundle 44, I think it's going to turn out to be volume 2, and it's going to be document 8 and it's going to be page 94. Is this what you're looking for?

A Yes, this is what I'm looking for.

Q So we see repeat, effectively, in the bed days----

A Yes.

Q -- in the right-hand column over three years.

A Yeah, and can you please scroll down so we see the data that are not imputed which starts----

Q So if we can move 2008, January, to the top of the screen?

A Yes, that would be helpful.

Q So, yeah, pull it up as far as you can.

A Yes, so we can see 2009 and so on.

Q Maybe look at the next page as well? Yeah, there we are, page 95. So what's the point you want to make?

A So there's sort of a standard range of bed days that we see in there, and it's imputed by the average per the month. You can see sort of fluctuation by months, but it looked like there was sort of a similar fluctuation throughout time. Imputing----

Q So that each month is broadly similar to the other months, so that the same month (inaudible)?

A Broadly. I mean, that's loose, but it's broadly similar, and so I thought, you know, this was a discussion we had. I said, "I can impute these or not."

Q Why didn't you put it in the HAD report?

A I had probably forgotten while we were putting that report together, and I apologise.

Q Because one of the things we will come to, and I'm not going to go to them now but I hope you'll remember this, is that in the linear progression best fit line that's in, I think, figure 22, there is a distinct downward slope in the rate of infections at Yorkhill. And in the chart that I asked you to calculate for Yorkhill from 2008 to 2015 for Yorkhill, we'll discuss what it means, but there seems not to have been a significant trend. Is that roughly right?

A We'll have to go to the chart.

Q Yes.

A I remember there were-- trends disappeared. I believe for that one I also-- I want to say I did a 2008 to 22, but I can't remember which one I did that for.

Q Well, let's just look at them because I think if I'm going to put something to you, I think I need to make sure you've seen the context.

A Thank you.

Q So if we go to page 118 of volume 1. So bundle 44, volume 1. We'll come back to this chart many times. Figure 22, right at the top of the page, and zoom in so you just see the rate.

A Yeah.

Q You've got a downward slope to that trend line for the Yorkhill data.

A Yeah.

Q Yeah, and there's no narrative in the text about it being significant, but it's there.

A Yes.

Q And it's downwards.

A Yes.

Q Right. If we hold that thought and we look at volume 7, 44, volume 7 and we go to-- fight our way through the document, and we go to page 57. This is the same data between 2008 and the move, am I right?

A That is correct. It's the same data, but I want to clarify that it's a different statistical model.

Q Thank you for that. What I'm wondering here is, to what extent is there a risk that the existence of those three years at the beginning is influencing the conclusion you reached in the HAD report that there was a downward trend at Yorkhill? We go back to volume 44-- that one, yes. Is there any risk that those first three years might influence that trend?

A I would say that the risk is slight. There's always a risk, right? And there's also the reverse risk that I want to mention that if you do not have enough data, then you don't have the power to see a trend that exists. But what you can see is the-- if you look at the peaks and troughs, so the way these models work is what they're trying to do is separate the noise in the data, which there is a lot of because we fluctuate a lot, from what's really happening in the data in terms of a trend or when we get to the GAM models in terms of the wiggly curve. So what you can see there is we don't have exceptionally high peaks at the very beginning.

Q But we have no zero.

A We have no zeros.

Q And so the reason I'm asking this is, we'll come back to the meaning later, but you didn't mention it. Why was it necessary to do at all? Why not just use the data you actually had, so eight to the move?

A In hindsight, I probably should have modelled both.

Q Because the length of the period from eight to the move is what, seven years? And the length of the period you are content to plot a line for in the children's hospital is seven years. So what's wrong with just having two sets of seven years?

A So, again, the more data you

have in a time series like this, where you're looking at data over time, the greater power you have to detect, to look at any changes, especially in the middle, and this becomes more relevant when we're talking about the GAM models, because all the data are together, and it's looking at this sort of separation here between 2015 and in the move.

So, throwing out data always creates a power problem, and that is an issue with a lot of the stuff we see here. As I said, in hindsight, it would probably be helpful to model both, which we end up doing. Of course, I move away from linear regression models which you see here.

Q Well, if you take that off the screen, before we look at the charts, there's just a question about process, because I want to clear up my understanding of something. You receive these two big-- well, not as big as this Inquiry one, but two big spreadsheets: one for adult, one for paediatrics. You carry out a deduplication process.

A Yes.

Q I'm assuming that's done in a software package.

A Yes.

Q It's not done manually or in spreadsheets.

A No.

Q And so if we go to the

questionnaire 2, which is bundle 44, volume 7, document 1, page 7, and we look at question 40, we asked you a question-- No, it's not that questionnaire. Sorry, I've referred to the wrong document. Give me a moment. If we go onto page 8-- Ah, that's where I've made a mistake. Volume 5, page 8. Page 7 of that.

Right. So, at 44, volume 5, at page 7, the Inquiry team looked at your tables 11A to 14 in the HAD Report. Before we actually had, I think, appreciated that you'd given us the environmental BSI count spreadsheet, we created a manual total, because I think I was just interested to see whether there was any comparison I could make with other epidemiology I'd seen. When we got the spreadsheet, we noticed there was a difference. So the spreadsheet you'd been provided in the spring of this year, after the request, was different, as this table suggests, by small amounts each month. So what I wanted to understand, if you could explain how it was that difference came about.

A Yes.

Q So, how did it come about?

A So, I write a code to analyse the data and clean the data and I reran that code. So I was-- The Inquiry team requested a bunch of chalked up datasets from me----

Q We did. Yes.

A -- and so I reran that code, and because I was working quite fast, there were a few items that I missed. So some of that was this deduplication with organisms that have to be grouped together, the E. coli, because they've got complex and they've got a few other things that didn't get grouped together.

Q So, you effectively had to redo your deduplication?

A Yes.

Q Had you retained the code from before?

A Yes. Yes, I just-- So, with that code, and there was also something missing which I added in later, and this is why I was saying as well with the-- it's a little bit different because there is hand counting work done with the tables for the clustering, which I moved away from, is that there's-- you know, there were items added in later and I had different versions of the code. So I had run----

Q So, eventually, the stuff you gave us in what was effectively I think part of volume 2 and we looked at briefly, that was different.

A Yes, it was different and it----

Q And it was different because of coding and changes.

A Yeah. And then I provided you that whole list again with this response.

Q Yes, and so the charts in the HAD response document and the short ones I asked you to look at, they used, effectively, a third output.

A The corrected, yeah.

Q So the first output creates the HAD charts.

A Yes.

Q The second output creates the stuff that you provided in volume 2, which turns out not to match the HAD data, you realise, and you provide us the third set.

A That's right.

Q How would you respond to the suggestion that making such an error was misleading the Inquiry?

A I would tell you that that was not my intention in any way, shape or form and I can only apologise. I was working incredibly rapidly, as you are aware. We didn't feel like we had enough time to get everything done.

Q No, you didn't. You had to turn it around.

A And so I just didn't-- you know, I didn't have the time to double-check all those things against the other items.

Q Is there anything unusual-- I mean, I can ask you questions about not disclosing material, and I have. So you didn't include the occupied bed day data in your original report. You didn't include the tables that show the incidence of each infection by month in the original report. Now, as a lawyer, I can say, "Well, that's not consistent with what you

said in your declaration."

A Right.

Q But as an epidemiologist, what's the practice there? Do you expect to include all this stuff?

A No, you do not provide any of that stuff. You provide the summary data, so that people can understand the data. Raw data is considered not very understandable.

Q Now, we received, from various people, submissions based on what the original report says, and no doubt we'll have submissions in evidence about what the second version of it has. The middle report caused NSS to say various things they said in their document, and the CNR to say things, and Dr Mumford and Mr Mookerjee to say things.

Is there anything wrong with effectively what you've done? Are you stepping outside the bounds of a normal epidemiologist by not including this stuff, making, what could be described as, these errors?

A So, no, I would say that that is not the case. So, in normal epidemiology, we produce our reports. These types of errors that are small are considered wobble. If it had been made clear from the CLO that we should provide datasets and provide code and provide everything else, we would have been happy to do that. So it is very normal in epidemiological practice to make datasets and code available at request.

Q Now, what I want to do now is to step into the charts. Now, just to explain to you what I'm going to do and what's the benefit for his Lordship and our other colleagues in the room, what I'm proposing to do is to start with the HAD Report, head to the end of chapter 7.2, after the clustering section, walk through the various charts, but not the conclusion bullet points.

Then go to the response document, look at that, look at some of the commentary at the time, but not resolve any inconsistencies, and then work our way through the ones you produced at my request. When we do that, pick up any anxieties you have about the questions I asked, and that'll be lunchtime, and then after lunch I'll try and ask you what it all means. Then if we still keep going, we'll do Aspergillus at three o'clock in the afternoon. That's the plan. Okay?

A Okay.

Q So, let's go to HAD Report, which is bundle 44, volume 1, page 115. You have a chart at Figure 19. Now, one of the slight problems of reading the report is that the charts don't always appear in exactly the right place

compared to the text they relate to. So is the text for this chart, following below it, in the page?

A It will-- In the paragraph that discusses it, it should reference the (inaudible 02:40.21).

Q Or is it actually back----

A No, that is the clustering there.

Q It goes back to clustering. So, the reason I ask is, if we look at Figure 19, and over the page onto page 116-- In fact, let's get the beginning of the sentence. So we'll go back to page 115. It's a long sentence. So five lines from the bottom. This is just seemingly discussing Figure 20, this paragraph. It says:

> "We can observe two exceptions to this, where in late 2006/early 2007 … there appears to be a more significant reduction in environmentally relevant organism than there are not environmentally attributable and again in 2015/early 2016 when the patients are moved from Yorkhill to QEUH, however, this does follow the pattern that is … observed directly before the move in early 2015. This type of difference in BSI pattern could be attributable to a wide range of factors, including changes to antimicrobial policy, capacity of laboratory services…"

Is that a counterfactual we ignored in the beginning, or do you want to add it to your list?

A Yeah, I think so, along with-- I mean, it starts getting mixed in with line handling, but contamination of the sample, right? When we're talking about non-environmental microorganisms, you can easily contaminate that sample.

Q

> "Policies on testing for BSI … natural patterns of microorganism attributable to environmental facts, as well as many others. In historical data without an extensive volume of additional information, it would be difficult to determine a likely cause."

Now, we're going to come back to all of that after lunch. And then it says:

> "While it is unclear what may be driving these two lower troughs in environmentally relevant BSIs compared to other types of BSIs at these time points, it is unclear that environmentally relevant BSIs largely follow the same pattern as those that cannot be attributable to the environment over time…"

Now, I'm assuming this is based on looking at Figure 20, which is below.

A It is based on Figure 20.

Q But I suppose if the troughs are in the environmental lines, which are blue and red, we'll also see them on Figure 19, without the non-environmentals there as well. If we go back to Figure 19 on page 115, now, two questions arise. Is the trough you have in mind the two zero months around 2007 on Figure 19?

A Yeah, that is correct.

Q Yes, and then the other trough is the slightly more than two zero months in late '15.

A Yes, when they move.

Q When they move. Now, two questions is, what, if anything, should I have as a form of concern about the fact that that first trough sits in the period when you have imputed the denominator? Does that cause me any reason to be nervous about relying on that trough as a thing?

A So, I am not nervous because if we look at this dataset-- So that was a comparison, right, between the environmentally attributable and the non-environmentally attributable. You can see more of those troughs in the Yorkhill data, for example, in 2000-- Go ahead and look at that 2012, month one. There's some there.

Q Yes, there's another one in 2009.

A Yeah, and there's one in 2013. You can see them. So the data from 2005 through 2008 largely look similar, troughs and peaks, to the data throughout. You have bigger peaks in 2008 and over in 2012/'13, but----

Q You're shrugging. We're just saying at this point, because the transcript really can't see shrugs.

A Sorry?

Q The transcriber can't see shrugs.

A Oh, I see.

Q Yes, so when you shrugged theatrically at the point you said those two things, I thought I should say it.

A Yes, yes. There's-- You know, it looks very similar.

Q So the point that's been suggested by NSS, Ms Cairns particularly, is that another feature of this chart-- It's worth saying I'm assuming blue is not BMT, it's actually-- This is paediatric data.

A Yes, this is paediatric data. I apologise.

Q Blue is RHC and North is Yorkhill.

A Mm-hmm.

Q She observes, and others do too, that in Yorkhill there are repeatedly months with a zero, and in the new

children's hospital, there are months with a zero, and then there was a long period of nearly, but not quite, three years when there are no zeros at all. Then the zeros, after a little flurry, start again. And the absence of a zero-- Because you'd expect zeros, wouldn't you, if it was random?

A Yes.

Q Is there anything we can take from the fact that the only place where there are no zeros is in RHC from, what is it, mid-'16, possibly second quarter '16, to towards the end of '18? Is that significant in any way?

A So, with the data that you're looking at here, there's no test for significance or anything else, and when you split the data set, and this is going to become important later, and compare it using a linear regression model, that does not come up or appear as significant when you use the GAM model, as we'll look at later. You can clearly see that there is something going on there.

Q I think the point that was being made by Ms Cairns was in fact simpler than that. It was simply if you have a random variation, you will always get zeros because that's the nature-- often get zeros. But the absence of zeros is something that we should ask ourselves, "Why is there an absence of zeros for the best part of three years?" Is that a question that we should even be asking?

A I think it's a good question. I think it's a fair question.

Q Well, we'll come back to that after lunch. What I want to do is to go back to figure 20. Now, because of course, figure 20 is quite hard to read-- I noticed you then produce a non-environmental chart, and that is on page 117, figure 21. Now, it doesn't say in this report whether these two best fit lines are significant. Are they?

A The first one is not, and the second one is, as I recall, if I'm recalling correctly.

Q But why doesn't it say it?

A I apologise for that. I don't know why it doesn't say it.

Q Because in fact one the oddities of this inquiry is it's not actually until----

A Later.

Q Well, you do it later on in this report. There's a mention of statistical significance and, perhaps prompted by that, we start asking, and you and Mr Mookerjee have both in different ways attempted to understand the statistical significance of reported numbers, but there's nothing in here that reports numbers.

A No, no. And I think this was largely because there was a lot of interest. Later on, I'm guided by myself.

Here I was guided by other people in demonstrating visually and testing less to keep things simple. I decided not to do that anymore, and----

Q Well, if we go on to the next page, page 118, we get to figure 22, which we will come back to. We've been to it before. Are either of these two trend lines significant? Linear regression lines?

A I would have to go back and check that.

Q Because the question that was raised in the NSS response by Ms Cairns was that she felt that there was something that might be going on in the blue line between '16 and '19 that required further investigation.

A That is correct.

Q And that's correct. You would agree with that?

A I would agree with that, and I followed up and did that.

Q Right. Now, what I'm not going to do now is look at the bullet points that follow at the bottom of page 118, 119. I want to come back to this this afternoon because I think they fall into the category of "What's going on here?", rather than the data. What I propose to do now is to move on to the response document. So, this is volume 5 of bundle 44, and I think we're starting on page 42. Now, I think this may be a simple case of mislabelling, but it possibly needs checking. Somebody raised this. So, figure 2E.3 is described as "Monthly incidence rates of BSI attributed to environmental microorganisms". Could it actually be non-environmental microorganisms and be mislabelled?

A This one that we're looking at here?

Q No, page 43. Sorry, this chart here. The reason I say that is if we go back to 44----

A There's no fear of----

Q -- volume 1, page 117.

A Yeah.

Q Page 117? Yes.

A Yeah.

Q So go back to volume 5, page 43. Could that simply be a mislabel?

A It could. It could be a mislabel. It certainly looks like the non-environmentals, doesn't it?

Q But in any event, this doesn't have regression analysis.

A No.

Q It's just a plot.

A It's just a plot.

Q And, in fact, the chart that was in a sense most supposed to be here should effectively be figure 22 from the previous document.

A That is correct.

Q Adjusted by the small adjustments that you say you made.

A That is correct.

Q Right. And so if we, staying with the paediatrics and the environmentally relevant, go to page 50, to page(sic) 2F.3. Now, what I want to do here is to understand firstly what the spots are and connect it to what we've seen before, and then understand why you did it and what you take from it, but not actually resolve any questions, or resolve them at the end. So the spots, are they in theory, with some adjustment the same as figure 22?

A Yes, they should be data.

Q Right. And they're monthly totals?

A Yes.

Q And in fact----

A Not totals, rates.

Q Rates, sorry.

A For 1,000 bed days, yeah.

Q And to return to our previous conversation, we have the zero months along the bottom. We can see them and we can see the gap I've just discussed with you.

A Yes.

Q Right. Okay. So what prompted you to try a different sort of best fit line?

That was Ms Cairns's comment about how linear trend lines may not show what we want to see, and I agreed with her and thought about this quite a bit more, which is pretty standard in epidemiology: you show people something, you talk about it and they come up with an idea that's very compelling. So I went ahead and produced GAMs. Do you want me to describe----

Yes, I'd like to try and understand GAMs by reference to a feature on this chart, and maybe it'd be efficient in time to do both at the same time.

A Okay.

Q So when you look at-- So what's the blue line on the chart?

A Okay, so a GAM is a generalised additive model, so you're effectively adding models onto each other here.

Q So what are the models you are using at the linear stage?

A And so a blue line is a generalised linear type model. So remember, we have multiple models in here that are additive. So when, for example, the red line and the blue line match, it does actually collapse down to a generalised linear model. So what you have is a linear component to your data, which exists in this type of data.

Q So is there a trend?

A Is there a trend, is it going up or down?

Q Right.

A Then you have what was called the smooth line here, and I kid you

not, the technical term in statistics is a wiggly line that you fit, and the wiggly line tells us if that trend is followed in a linear way or if it's followed in a flexible fashion. And then you have a third line, which I haven't shown here but you might want to talk about for completeness, is a seasonal trend line, and that just makes a mess over the top.

Q Yes, because it's every 12 months.

A Yeah.

Q Now, the colours, the shading, what are they showing us? Start with the blue line shading.

A So the blue line shading are the 95 per cent confidence intervals around the dark line, which is the fitted point estimate for every single one of those lines.

What does it mean to say that something has had a 95 per cent confidence interval, assuming for the purposes of this answer that we have no statistical knowledge at all?

So, when we fit a point estimate, we're saying we believe this is the estimate based on the data, right? We believe this is the estimate based on the data but – there's always a but – with 95 per cent confidence that could fall between these two lines. So if you look at the outer edge of our confidence interval, the true point estimate, because we're sampling, we're not taking the entire population, could fall between those outer bounds.

Q So the true position----

A Yes.

Q -- if you were to pick it at any point in that period is 95 per cent of the time going to fall inside those two lines?

A Yes.

Q And is 95 per cent a conventional assessment of sufficiency?

A Yes. Yeah.

Q Right. Does it work similarly for the pink shading?

A Yes, it does.

Q Right. Now, because it may become relevant later, is there anything important we take from areas where the two shadings overlap?

A Yes, so what that's saying-- And particularly this becomes important when the pink overlaps entirely with the blue, right? What that's saying is both of those components are telling you the exact same thing is happening at that point in time. So when we look at the data up until 2014, I would say, both those lines are following each other, which means we're seeing the same trend.

Q It's a linear trend.

A It's a linear trend.

Q Right.

A And then we see a deviation

from the linear trend, and when we see that, whilst the linear trend is still valid, so if that was significantly decreasing, it's still significantly decreasing. However, we believe the pink line or the red line, that it actually went down, came back up and then decreased.

Q And just at the end on the right hand side after 2022, is there any point about the overlap there? Does that tell us that we're back to a linear trend potentially after 2022?

A Yes. Yeah, they're back to linear trend, but you can see that red line starting to dip down. I don't know what would happen if we had more data.

Q And this is the question that I don't understand at all, and I'm trying to get an answer, and I'll try and answer it. You've observed in the following paragraphs that the dip that happens where the red line goes below the blue line, where the bottom of the dip-- And we might zoom in a bit just to make it easier. There we are. The bottom of dip is what, mid-'15-ish?

A Mm-hmm.

Q That dip starts in early '14.

A Yeah.

Q And after lunch, I'm quite keen to get an answer to the question I'm about to ask, but you might want to discuss whether it's a good question now. I suppose one reason the dip starts in 2014 is because something happens in 2014 that causes a change. That's one possibility. Would you accept that?

A Yes, I would accept that.

Is the other possibility that effectively the line is trying to get down to a point when there are lots of zeros and starting early to smooth out?

That is also a possibility. So we are letting the data dictate what is happening, and after lunch I can talk about, sort of-- the way the model is set up is we want it to be responsive to rapid changes, this curve, so we've chosen features that allow it to be responsive to rapid changes, and it could just be that set of zeros that are pulling it down.

Q And we might look at other things that you've done that might help us on that territory.

A Yes.

Q But we'll do that once we get-- Now, we come on to the peak. Mr Mookerjee hated it whenever I used the word "peak". What's your attitude to the word "peak"?

A I'm all right with that.

Q You're all right with the word "peak". Okay. So, obviously a peak involves a change of direction in trend. From just looking at this chart, are we able to tell when that change of direction happens, or should we be nervous about temporally fixing it at what looks like just

after Christmas 2017?

A So, what I would say about that, and no confidence intervals around it, we have a rough idea of when that is happening, so it's not happening in 2014 or 2015. It's definitely happening sometime in 2016, but to say, "Well, I could go through and count, you know what month that is. That is the accurate month", I'm going to be a little bit more relaxed.

Q You might think you have, what, a year or more of a possible period when it's changing, but we're really being very vague here.

A I would say-- No, I would definitely say that's changing in early to mid-2016. So it's not-- I would-- I would-

---

Q The change to go down again?

A Oh, I'm sorry, going up, and then we're talking about going down now.

Q Yes. When's it going down?

A So sometime-- It looks like from this-- But you'll see I did a change point analysis----

Q Yes, we're going to cover that -

A -- as well, because modelling data in multiple ways, if it's showing you the same thing, that really helps confirm what you're seeing. So here the peak is at January 2018, and the downward trend happens that directly under that.

Q Right. Now, what I think I want to do now is to look at the non-environmental bacteria, just as an aside, so that we pick it up in order. So page 51. So just in a sense to put this in context, this is in fact the same data as page 43.

A That is correct.

Q Albeit, 43 has got a different label on it. Right, okay. So you've discussed what the lines are trying to do. What do you see in terms of points when things change in this chart?

A Yeah, so I guess the other thing that I should add for completeness is, each of these lines has a statistical significance value on it, and if that red line – because we'll see this later on – is wiggly, but it is not statistically significant, you shouldn't believe the red line.

Q Yes.

A So I just want to make that clear----

Q So what about these two lines?

A Yes----

Q Are either of them statistically significant?

A The-- In the previous graph and in this one, they're all statistically significant. I just want to make that clear.

Q When is the point when we have an upward and when is the point we have a downward tend?

A Yeah, so just before January

2014 this starts going up, and it peaks what looks like maybe March 2017. Roughly. It's just after that. January 2017, the line, and then following that it starts dropping.

Q Now, I should have asked a question about the previous chart. I asked you about the drop below the line, whether it's the chart trying to reach a particular (inaudible - 12:38.13) of zeros, but I suppose it's harder. There's a less obvious little group in the corner here, but is the upward trend in this model attempting to get up there to catch the tops, and is the dip down happening because it's desperately trying to get down to the bottom right-hand corner by the end? Is there any of that going on in the model?

A So-- the way these models work is, they're smoothing that information at the top, and remember I said that they're trying to account for noise. Right? So the more noise they can see over time, the better they can account for it, because they can say how much noise is real. So it is going up because you see, sort of-- If you look at your points in that time period, right, you see very little below the blue line and you see quite a few above the red line. So that is a very smooth pulling it up. The-- The lower end that you're talking about----

Q In '22?

A Yeah, in '22, also looks pretty reasonable to me. There aren't a lot of zeros there and there is a lot of data falling down there, so I-- I feel like there is a lot of data showing the same pattern in these-- in these areas.

Q Remember I asked the question about trying to locate the change points in time----

A Yes.

Q -- and you suggested it's somewhere in the confidence interval shading?

A Yeah.

Q Now, I suppose one could ask you about the peak at the beginning of '17. Where are you comfortable in describing that peak happening at the point when it begins to go down again?

A So, I guess what I would do is-- If we want to look at the change point analysis, I would think about-- it could from where the change point analysis sets it, or it could be from where this sets it, and there----

Q Well, we'll do both. So from this one----

A Okay. From this one I would say that you've got your peak, and I'm going to call that probably March – it might be February; I have it listed on here because I had the programme provide that – is where you hit your peak, and

everything after the peak is down.

Q Right. I think you said March----

A March? Perfect.

Q -- on page 52. The rise point-- Now, you didn't I think say in the report-- You have August '13 as the rise point.

A That looks correct from what I'm looking at here.

Q I may have misunderstood something you were saying about the confidence interval shading, so I want to just see: does the fact that the smooth line takes a rather gentle rise compared to the previous one-- less abrupt, I suppose, the better word, does that cause the period when the change might be taking place to sort of widen in time from left to right across the chart? Am I imagining that?

A No, it is wider.

Q Yes.

A Yeah.

Q So what sort of a range of dates for that upward peak-- I mean, August is obviously the centre of it, for that's what you've written in your report, but what's the sort of range of dates----

THE CHAIR: Sorry, just clarify there: you accepted from Mr Mackintosh that looking at the upward change, it is wider because it's a less steep curve.

A That is correct.

THE CHAIR: Did you accept what I----

MR MACKINTOSH: I'm just double checking that, my Lord.

THE CHAIR: Okay, right. I'll----

MR MACKINTOSH: So, if it's wider in time, potentially, and you have in the report identified August '13 as the upward point, does this have a range of possible points, or is it really just August '13? From this chart.

A From this chart, it is August '13. What I want to be careful about is what reality is. Right? So, based off of actually any of the statistical models, I am not going to claim that reality says it's that exact point, and that's because statistical models work with the data with-- in the ways that they do with all their underlying assumptions and the ways they move. So, you know, was that pulled up slightly later? Probably not slightly earlier. Looking at the data, maybe.

Q Well, let's look at your attempt to see what the data told you. So on page 52, the whole page for this one, you tell us in 2F.18 that you attempted an exercise-- Well, what was this exercise you attempted that you've illustrate in 2F?

A Okay, so this is-- This is the-- This is a different way to model what we were looking at before, and you want to use multiple models because those give you more confidence in what you're seeing. If you use two different modelling

techniques and they give you two different answers, you should be very suspicious about your interpretation of the data. So what this is is a change point analysis model. It is a-- The other ones were negative binomial models, that is the family.

So when we fit a general linear model, we have to specify families and links, and-- a negative binomial deals with lots of zeros so that they are not causing a problem in your model-- or lots of extreme events, I should say. This is a Poisson, which is a standard statistical model for count data. Negative binomial is a form of a Poisson model, but it's a special form of Poisson model. So I can't use the negative binomial model to do the change point, I have to use the Poisson and it just doesn't work in there.

Q Now, you've identified some change points here.

A Yes. So----

Q Before we look at them----

A Yeah.

Q -- how significant are they?

A So these are all significant change points. So when a-- In a change point model, when the model calls a change point, that is significant. So what it does-- This is a very different type of model. It says, "I can call this-- this incidence rate the same, the same, the same," right? "Given all this variation, I can still keep calling it the same," and at the point it can't do that anymore----

Q It changes?

A -- it makes a significant change.

Q So what you have is you have a rate of 5.5 until some time in late '10, steps down to 2.8 until sometime in late '16, steps up to 6 until some point in late '19, and steps down to 1.8?

A That's correct, and those are the average, right?

Q Yes.

A Yeah.

Q I just wondered whether, if we sort of flick back to page 50-- I remember your evidence was that in the first period to sort of (inaudible – 12:45.32)-- I can't remember, I think you might have said '13, it's a linear trend down.

A Mm-hmm.

Q Then we have this feature that we've seen. If we go back to 52----

A And you see that there.

Q Effectively, can we see something similar happening here?

A Absolutely. So you could-- And actually, you know, I-- It's just good as an illustration. I-- don't want to reopen this can of worms and talk about it, but Sean MacBride's analysis also did a change point analysis, and you see-- and he actually, like, attached them, and you can see that same trend that you see

on his decomposed trend line, which is what we're looking at here.

Q In terms of the dotted lines that show the chain steps, should they perhaps have confidence intervals on them? Or am I missing my metaphors here?

A You can't get confidence intervals on these.

Q Okay. Seemed a nice idea. If we go over the page to 53, you did the same thing for the non-environmental.

A I did.

Q Now, if we flick back to 51, just for the purposes of having it in our minds, do you see any connection between these two analyses?

A Yeah, it looks the same to me.

Q How would you describe what you're seeing in the two, now that you've seen both?

A So we see this downward trend to begin with until, you know, this time period in late 2013. It goes up, you have this peak, it comes down, and then it drops below that-- that blue trend line, and-- you sort of see that when you-- if you flip back over.

Q So just to sort of recap where we got to at the end of this section, I'm going to put something to you which I think is summarising your position, but I want you to tell me if I've got it wrong. That, in respect to the environmental, there's a linear downward trend at Yorkhill, and then there's a peak in early '18, and there's something happening between the two which may involve the rate starting in-- drop starting in '14, or that may be a feature of the GAM. We have to look further to find out more. Is that broadly----

A So, sorry, I lost you when you said something about dropping in 2014.

Q Let's go back to page 50.

A Are you talking about the dip, the early dip?

Q Yes.

A Yes.

Q Basically you're saying, from the beginning of that chart to '13 we've got a linear downward trend – not very steep, but it's there – and then there's a feature that involves a peak in early '18, and the join between the two seems to involve a dip starting in early '14, and there's two possible things going on. One is, actually it is starting then, and the other is the data is actually just creating a smooth, and we have to think a bit more about that.

A Yeah, and then-- and then what I would add to that-- I would agree with what you're saying, and what I would add to that is, all those zeros-- that line is just following what it sees in the data. There are a lot of zeros there.

Q There are a lot of zeros in that.

A Yeah.

Q Then if we do the environmental one on the next page, that's page 51-- (After a pause) There we are. No, page 51. Yes, I think you were there already, that's why. In what way is the pattern similar or different that we're seeing between these----

A Yeah, so-- so we clearly see the rise in infections happening earlier.

Q Yes.

A So they're happening already at Yorkhill.

Q Yes.

A And they're continuing on when they move the children, and they're coming down earlier as well.

Q Right. Okay.

A But persisting for longer. That peak is tighter.

Q It's a flatter-- Well, not flatter, but it's a wider peak.

A That's right.

Q Right. What I think I need to do is decide-- I think I'm going to pass over clusters for a moment----

A Okay.

Q -- because I think we're going to come back later, and do some of what I've described as your joint work with Mr Mookerjee. Now, what we decided to do as the Inquiry was to ask the two of you to meet. I don't want to have the two of you discuss your conversations, because the question was, "Could you agree?"

A Mm-hmm.

Q And you actually didn't agree, ultimately, on a single report. That might have been down to the lack of time. So I wanted to look at the two things you produced from that conversation. I'm working on the basis that the conversation was helpful, but you're not going to tell me how. So a sort of----

A That's fine. Yes.

Q -- (inaudible – 12:50.13). So if we go to volume 7 and we go to document 4, page 51, you helpfully produced a lot of this report, and it contains-- firstly I recklessly asked you the Aspergillus questions first and, secondly, you included a lot of calculations, I think possibly having been challenged about not doing it before.

A Yes.

Q Yes. So let's pass over Aspergillus and go to page 57. Now, I think this is the point when I ask you to express your views on the very nature of the exercise I asked you to do. So I asked you to, effectively, attempt some form of trend analysis for four time periods.

A Mm-hmm.

Q So that was Yorkhill from eight to the move?

A Yes.

Q RHC from the move to the

date we suggested, which was the decant from 26 September 2019 which some people seem to think is important, and from that date until the opening of the new Schiehallion, and then nervous those periods might be too short, the whole period of the children's hospital from the move to the opening of the new Schiehallion.

A Yeah.

Q Now, do you have any concerns or are there positives and negatives of doing this exercise?

A Yeah. So I think I'll start with the negatives and why I didn't do this in the first place. So the first negative is that when you start cutting up the data, you have decided when the change point should occur. You're not letting the data tell you when it's occurring. So when you do that, you're imposing sort of an a priori belief on the model and you're preventing the testing of something happening in there and I show that in the later models where we test that.

So, I can sort of say, "Yeah, I can compare that this is-- let's just make up a hypothetical situation. I see a downward trend in this chunk. I see an upward trend in that chunk," and you're hypothetically imposing that they have a relationship. When you keep the data together, you can actually test that relationship. So when you cut it apart, you impose changes that may or may not be real, and you also eliminate your ability to do high quality testing on it. The third thing you do is you reduce your sample size. So, remember I said the longer time period we have, the more the model can see the noise and separate the noise from the reality, and that's actually what we want the model to do.

The other thing that happens is there's less power on the ends because we don't know what's gonna happen in the future or we can't see what happened further in the past, so it has a harder time picking where the right spot is on the ends, and so your power comes in the middle. So if you cut it right where everyone moves, you can't actually see the dynamics that occurred when they moved.

Q Are there any positives?

THE CHAIR: Just before we move to the positives, I wonder if I missed the second negative. Your first negative was you cut up the data, your (inaudible) investigators----

A Are imposing the belief of where the change is.

THE CHAIR: No, it was the point that comes after that.

A Is that you can no longer test for that like a sudden jump or something in the data where you can if you leave the data intact. Right? You're just sort of

comparing two things which isn't a proper test and you can say, "I group this and I group that," and you can put a test on it, but that's not really a proper test. That's not what we want to do in statistics. We would rather test it as one unit. And then the third one is that we lost our power and there are two issues about losing power in there.

MR MACKINTOSH: Given your concerns, and if we go to page 118 of volume 1, why did you put trend lines on the two hospitals, on 118?

A So this has been an evolution for me, obviously, in working on this report and I modelled this very differently, again, here. I was saying, well, this is a pretty standard way of modelling things. You saw Mr Mookerjee did this as well, you know, use a linear trend and compare those linear-- you know, just show-- that visually show and that those trends are different, right? It isn't as good as I work through this and think about things as doing it the other way. You can do that. It is done all the time.

Q And you did it.

A It is not considered-- yeah, it's not considered incorrect. But when Ms Cairns had addressed, "Is that the best model?" I thought about that for And they said, "Actually, no, I think there's a better model," and now I'm convinced that there is a better way to do it.

Q I suppose the reason I asked you to do it was because you produced a report with a-- and we'll get to what it all means, but the question that I had in my mind was, look at the environmental figure on page 50 of volume 7, it had this dip turn before what I know to be the move.

A Mm-hmm.

Q And so I wondered if you'd accept the proposition that, whilst you may not know whether the move to a different water system caused an increase or a decrease, it's not an unreasonable question to ask, was there a change?

A No. So these are not unreasonable questions, and it's not unreasonable to cut up the data. I'm not saying that you asked an unreasonable question, and I was happy to do it, but as I thought about it and looked at the data, I felt that we lost information by doing that.

Q I just wonder whether, and we'll come to what the cutups show, whether the exercise of doing the cutups helps answer that question about whether that turning point before in early 14 is a real thing happening then, or an artifact of the model trying to deal with something that actually happened, and this is the hypothesis that we're testing on the move. So whilst the principal technique is the one you've done in the response

document, do these charts help us explore that issue, which that chart can't do to the same degree?

A I'm not sure I would agree that the other chart can't do it. I would agree that-- I'd always agree that modelling things in different ways and exploring what happens there is very helpful to our understanding because, remember, we're dealing with retrospective data, this is epidemiology, we're not there, we're not controlling things and so this is a very hard problem to solve. A very hard question to answer.

Q It's just that one of the problems, I suppose, to mention is if we go to the appendix of Dr Agrawal's instruction letter, so that's this volume, page 242, the question you're asked is about the widespread contamination of the water system. Whilst I appreciate you didn't read the HPS report at the time describing what it is, no one is suggesting that contamination happened at Yorkhill. So in a sense, one of the two key questions that seem important from this question is, did the move to this new water system affect the experience of the patients as recorded in their data? Infection rates. So when, after lunch, we go and look at the charts that you created and chop up, what I want to suggest to you is that whilst we should be careful, are they of assistance in the journey we are now on?

A Yeah.

Q Right. I think, my Lord, this might be a good place to stop before we turn to the other charts after lunch.

THE CHAIR: Yes, we'll take our lunch break. As I said, Dr Drumright, we usually sit again at two o'clock, so if I could ask to be back then and I hope someone will help you with lunch.

A Thank you.

(Adjourned for a short time)

THE CHAIR: Good afternoon, Dr Drumright. Mr Mackintosh.

MR MACKINTOSH: Thank you, my Lord. Now, Dr Drumright, where we got to before lunch, I think, was we were discussing the merits and demerits, or the demerits and merits, as it were, of cutting up the data to look at shorter sections. It occurred to me we should just look at those exercises. I wonder if we can go to bundle 44, volume 7, page 57, which should be-- Has she looked at 56? 56 is the question I asked, "What was the overall BSI incident rate attributed to environmental microorganisms among paediatric and haemato-oncology patients in Yorkhill for the whole period, January '08 to May '15?" And then during the same period, "What are the GAM linear plus smooth components against monthly incidence rates?" I think

you've followed your own route for the second half to some degree, but we'll get there in a moment.

So, just in terms of calculating the rate for the period, I've asked you to do this for all of these periods, and I asked Mr Mookerjee to do it too, and the answers aren't particularly different in terms of numerical value. Now, again, remember the-- Would you agree with that?

A I would agree with that, yes.

Q But is it an exercise that's worth doing, calculating rates for relatively long periods in order to see if artificially putting a change point in has a change?

A So, I would argue that's two different things. I'm going to split that up. So the change point-- putting the change point in, I'm happy both to let the data put that in and for us to explore that, that is slightly different than aggregating all of your data over a long period of time. So when you do that, right, we are interested in peaks and troughs, and we're interested in what is noise and what is real in those peaks and troughs. So when you collapse all your data over that time period, you just lost all that information.

Q I just wondered if it was in any way related to the output in volume 5 of bundle 44, on page, say, 52, where I

appreciate the data has told you the change points, but it has also collapsed between the change points. So whilst you'd have to be aware that in the other attempt, we impose the change points, or I impose the change points, can we gain any information by looking at that landscape of rates between change points and the imposed landscape of change points, and thinking, "Are they similar, are they different? Is it telling us something?"

A I mean, I understand what you're doing now when I look at that. I think there are nuanced differences where this is a slightly statistically more palatable way to do it with the change point, because it's taking into account the noise. It's looking over time, and then it's giving you an estimate for that time period. What we are doing is just collapsing data down, so it's not quite the same thing but, you know, I'm happy to go along with this as sort of an exercise of reasonable interest.

Q Well, what we'll do is, we'll move on to the chart itself. So back to volume 7, page 57. Now, you narrated on 57 what you see in the chart below, but perhaps you could just explain to us – albeit I set the period – what this chart may or may not show that is significant.

A Yes. So, again, these are the exact same models, the GAMs that I

used previously.

Q And they're the same as the long chart, 2F2.

A Exactly. I didn't change the model, I just changed the time period and, of course, because it takes less time and this was a much more rapid exercise, I didn't put in those confidence bounds on that line. And if I'm just revisiting, the p-value for that linear trend was not quite statistically significant, is what we're looking at there.

Q Because if it was essentially 0.07, it would be----

A 0.05 or less.

Q Yes, it comes out to 0.07.

A Yeah, so, you know, this is a little bit-- this is where we're kind of on the edge, right? So, when you study to-- Statistics people say, "Surely 0.06 is as good as 0.05," right? That's an arbitrary cutoff that we've set in the world of statistics. So something between 0.1 and 0.05, it's not statistically significant, but you might think that's interesting.

Q And does it have any connection to that downward trend linear component that ends in 2013, that you found in the earlier chart?

A It does, and so what I might say, looking at this, is, we chopped up the data and that took away our power. This is probably a linear significant trend because that is what we see when we have the full power of our dataset.

Q Right. So, if we go on to the next page where the same question was asked but now from move to decant. Now, again, you calculated the rate, that's 4.7 as opposed to 4 and a bit, so it's not much different really, is it?

A No.

Q But, of course, there's a big difference in the time. Would you agree with that, from the beginning to the end here?

A Oh, absolutely, and I believe both of those are statistically significant. I'm just double-checking. Both of those lines are statistically significant, and there is a very big change. It goes from very, very low to fairly high. And what I would trust is that red line, as you're beginning and ending, not necessarily that blue line, and that's because that red line is our smooth line, and we know that's statistically significant.

Q Right. Because it's statistically significant on the other chart?

A It is on the other chart as well. Yeah, yeah.

Q Yes, exactly. And how much weight should we take from the impression that might be here, that there is a steeper growth in early '16 than later on? Is that something that's of worth note, or is it just too fine to worry about?

A I would be careful of that

interpretation. I would just say what this is telling us is that there is a definite significant increase. There is some wobble in that increase. So I guess what you're saying about the steepness is you're going above that blue line and now you're going below it, right? You're saying it's coming up a little bit.

I would believe that trend, that is a smoothed trend. So, again, I would take it lightly as I would with most statistics on retrospective data analysis, but it certainly is telling us something, and it certainly is consistent with our change point analysis and it's consistent with our GAMs from the longer chart----

Q From the long chart, 2F5, right.

A And all that is good. Consistency is always good.

Q Yes. So, if we go to the next page, just to keep up the topic, we calculated the rate for this shorter period from September '18 to February '22. Now, I appreciate it's a much shorter period. So am I right to hear your anxiety about the period much more loudly at this point?

A Yes.

Q Yes. But the rate, for what it's worth, you calculated was 2.8 and that seems lower. Is there anything that we worry about? Is that interesting or just put it in the pocket and worry about it later?

A Yeah, I don't worry about that being lower. What I am slightly concerned about is that I think we're missing a downward trend that is happening in this time period, but----

Q And, in fact, the next page, when we asked you to do the whole period in the hospital, firstly, is this period a more useful length or----

A Yes, I would definitely say that.

Q I'm going to ignore the rate, because it's up and down, but what do you feel that this exercise is teaching us about the-- It teaches more than the long chart, 2F5 or not?

A So, I don't feel it does. I feel it's showing us a very similar pattern as the other one, and it's fairly consistent in even where we see the upward trend and the peak of the curve there. So I think that this is basically showing a scaled down version of----

Q Right. In fact, you put the chart on the next----

THE CHAIR: Sorry, just give me that again. Showing a scaled down----

A Version of the long chart with all of the data from 2005, all the way through to 2022.

MR MACKINTOSH: Which is 2F3. I've been calling it 2F5, but it's 2F3, and actually you've repeated it on page 61.

A Yes.

Q On the next page, please.

A Just to make the point.

Q Now, I do want to repeat the question I'm about to ask you.

THE CHAIR: With apologies, can I interrupt? Again, it's really just my noting. When we were looking at, I think it is, Figure 2.6, I noted you as saying it shows a downward trend, or you miss some of the trend. Now, did I----

MR MACKINTOSH: 59?

A Could we go back to 2.6?

Q Page 59, please. Thank you.

A Oh, so with this one, if I was reading up above, in 2.5.2, this line is not statistically significant for the linear trends. So what that is saying is that's essentially the same as a flat line even though we're seeing it go down. It could be the same as a flat line and, what I'm saying is, I don't necessarily believe that. I think it's still going down.

THE CHAIR: Right, so it----

A I think it's a power issue.

THE CHAIR: Sorry to be so slow but----

A It's all right. I think it's a power issue, and these are very difficult concepts, so don't worry. Please feel free to ask.

MR MACKINTOSH: Is it because it's a shorter period of time and----

A Because it's a shorter period of time, there are-- So the idea of power, the more observations you have, the greater the power. This period of time is really short compared to all the other ones and so we've lost a lot of power, and of course we have a lot of noise you can see in there, and it can't make a decision, the model, that there's a difference between noise and and trend.

Q Now, what I'd like to do now is to look at a chart that Mr Mookerjee calculated using, I think, your data or a very close proxy of it.

A Yeah, I think he used it. I sent over the data, as you recall.

Q Yes. So this chart is in his report at the same time. It's the same bundle. It's page 47. Now, he hasn't used GAMs – you're shaking your head – and he hasn't used negative binomials.

A No.

Q He's done linear trends.

A Yes.

Q But I wondered whether this exercise that he's carried out, in a sense adds anything to your understanding when viewed alongside the charts we've just been looking at?

A I would say no, and I would say no mostly because I'm being a bit statistically difficult. This is not just a linear model, like a general linear model. This is a linear regression and it does not meet the assumptions of a linear regression when you have that many zero time points. So if this was a GLM

where he was looking at at linear sections, it might be of greater interest, but I'm just concerned that the model is incorrect.

Q So in a sense the flaw in picking this linear regression model is because there are so many months for most of the period with zeros, it gets confused by that.

A Yes.

Q And the same would be true if it had lots of very high ones as well.

A Yes.

Q It's the extremes that are messing it up.

A Yes, and with linear regression zeros are a particular problem. So if you go back to the models, right, if we go to a generalised linear model, that's different because, as I mentioned, the families before, you get to pick the family which is telling you something about the underlying distribution of the data, and the underlying distribution of the data in a linear regression assumes a normal curve, and with counts, you don't have a----

Q Because time doesn't pass like normal curves.

A Yeah. No.

Q In a sense, in a normal curve, it would be like the population. You'd expect a normal curve.

A That's right. That's right.

Q But events happening over time, there's no reason to think they would follow a normal curve.

A That's right, and with something like this, thankfully, you have a lot of zeros.

Q Right, okay. So go back to page 61 of the same bundle, and I ask my plea for help. So if we go back to the evidence we just discussed before lunch, the start of the dip. So if my colleague can zoom in so that 14 to 16 is a bit bigger for the benefit of anyone watching this on the YouTube feed. Now, we discussed the bigger dip, the dip starting in early '14, and I think you accepted that it might be a real thing or it might be somehow an artefact of the process. Now, we can come and look at evidence from the real world later, but does what you've just done help you understand which of the two it is or are we still slightly mystified?

A Yeah, it doesn't help me understand that.

Q Okay.

A Because we basically put our cut right kind of in a critical spot for that dip, and remember I told you we lose power on the ends.

Q Right. So what I want to do is to take that off the screen and look at the non-environmental stuff that we did-- I asked you to do. I think I know what the

answer is going to be, but it's page 62 of the same bundle. So we simply asked you to cut at the change. Does the exercise that we carried out on this page, page 62, which shows a plot in the Yorkhill, and page 63, which shows a plot at the new hospital, help you understand any more than the previous chart?

A No.

Q No. If we go over to the next page, on page 64. Now, it's gone a bit sad, so we'll go back to volume 5, page-- I think 52 it is. Hopefully we'll see its happier friend there. I think it's page 52, at least I hope it is. Definitely not. Take that off the page, bundle 44 volume 5, 51. Yes. So we're back where we were before lunch. Right. Now, what I wanted to do is to go back to our conversation about the two scenarios that you were presented with and the counterfactuals, and if I remember where we got to, there's the contaminated water system scenario, which at the time you wrote the report you didn't know the dates for.

A Right.

Q There's a ventilation scenario which you didn't quite know the detail of, but you knew it involved air changes.

A That is correct.

Q We have the CLABSI counterfactual; we have the nursing single room counterfactual; we have the teams counterfactual; we have the antimicrobial resistance counterfactual; and maybe we have the lab capacity counterfactual. That's the list.

A That is the list.

Q Right, okay. I'm trying to think of a way of asking you a sequence of questions where we gradually introduce evidence from what actually was going on, if we can use that, outside the data; and either you can tell us what you think is going on or, if you are concerned that you don't actually have enough information because you've not been given access to early enough, you tell us what to look for.

A Okay.

Q So what I thought we could start with is to try and understand from your point of view what the mechanisms are of the counterfactuals and the scenario. So we'll start with the laboratory – take that completely off the screen, please – not working as efficiently. Now, if that was going to affect the rates of either the environmental or non-environmental groups, what would be the mechanism?

A So, this is the most complicated one, I think, in terms of what would be the mechanism. You could see a wide variety of issues, so you could imagine that, for example, the laboratory is contaminating the samples when they come up. So this is one problem you

could have, in which case you're going to largely see, but not entirely, non-environmentals, right? Because most people are contaminated, like their hair----

Q Because actually, the number of samples for non -environmentals is much higher.

A Yes. And so that could be, again, like I said-- And maybe we shouldn't call it a lab issue. Maybe we should just call it a contamination issue. So that could be happening with the nurses, it could be happening up at the lab, you know, whoever is drawing blood, whoever is handling the blood. That's one issue.

Q Just before you move off that, would that show any temporal change? I mean, if your lab system wasn't brilliant, wouldn't that just be the same over a long period of time?

A Yes, if everything stayed the same. But if you had problems like you were understaffed, for example, or your staff was in training or something like this, then you might see it for a period of time and it might go away.

Q Would we expect to have heard that from laboratory microbiologists and lab scientists who are giving evidence in this Inquiry?

A I would hope so but----

Q It depends who they are.

A Yeah. I can't give you any more information than that. I think that was not on the original counterfactual list because I think it's the most unlikely scenario.

Q Let's move on, and we'll not do this in too much detail, to antimicrobial resistance. What's the mechanism that antimicrobial resistance would affect the environmental group and/or the non-environmental group? How would the mechanism work in your understanding?

A So, I am going to ask you to get details on this from my colleague Dr Agrawal, but in general there will be antibiotics that can drive up cases of certain types of microorganisms, and a lot of those, as I understand it, are environmentals. And so if you start using more of those, prophylactically particularly, you end up seeing a whole rash of new cases in certain environmentals.

Q And would I be expecting to ask Dr Agrawal about how that counterfactual hypothesis connects to prescribing patterns or the number of microbials that were in the sample? That would be the connection.

A Yeah. I think he could probably look at the graph that I've produced, as well as the other information that you have.

Q And if we have a chart that

shows the incidence of meropenem-resistant bacteria over time, he'd be able to look at that, wouldn't he?

A Yes.

Q Yes.

THE CHAIR: Maybe this is just to state the obvious. If one was testing the hypothesis that the choice of antibiotic was affecting the total population of potentially pathogenic microorganisms, one way of at least taking your investigation further would be to find out what was actually being prescribed over particular periods in particular patient cohorts.

A Yes, that would really help.

MR MACKINTOSH: Okay. If we move-- And so other than that report that turned up in the folder which you hadn't seen before, have you been given any reports about prescribing patterns in the hospital?

A No, no.

Q Right. So then we move on to the team dynamics one.

A Yeah.

Q I mean, I recognise this is quite a "How long is a piece of string?" type question, but how would team dynamics affect these two classes of infection rates?

A Yes, so given that they both go up, and that's very important, I would expect to not see a huge difference in-- You know, I wouldn't expect environmental or non-environmental to go up and the other one stay flat with team dynamics. So just to make that clear, it's because they're both going up.

Q But they go up at slightly different times.

A Yeah, and that doesn't bother me so much because of course if you've got issues, care issues, you're likely to see your non-environmentals first because it's so much easier to get a bloodstream infection from a non-environmental. So everything that's on your skin-- You know, that's everything from inserting the line to managing the line. There's a whole host of microorganisms that are very good for us waiting there to become opportunists with a line. So you expect to see them go up and come down first. So if they have a problem you'd expect them to go up first, and if they manage that problem you'd expect them to come down first.

Q So we turn to-- At this point, the way you talk about it, I'm wondering whether the line issue and the nursing issue are actually the same issue?

A They could be the same issue, so patient safety issues related to dysfunctional teams, poor handover, disagreement which means that a policy doesn't get brought down to the group from one of the management positions,

other things like that. So these are the types of errors we see happen, particularly, as I said, with handover or with not implementing policies or that sort of thing, or people feeling like they don't need to because there's a discussion about this person's policy, "This isn't very good, I don't like this," so then an interpretation from the other ranks that I don't have to do this.

Q If people sort of don't do things consistently, they don't follow instructions.

A Yeah.

Q Did you read the Case Note Review Overview Report?

A Can you----

Q I'll put it on the screen. So it's bundle 6, document 38, page 975.

A Yeah, I don't believe this is one that you pointed me to, and I don't believe I've----

Q Well, it's the subject of the criticisms by the HAD report in chapter 4. So, you signed the report, so did you read the document you're criticising?

A So, from-- No, that was Professor Hawkey's commentary.

Q I appreciate that, Dr Drumright, but you signed a joint report in which you criticised the Case Note Review and said we should ignore it, and we had the Case Note Review authors provide a rebuttal document----

A Which I did read.

Q Yes, and they took a lot of time to do that, and no doubt that cost time and money. And I need to ask you about chapter 8 of this report. Did you read it?

A No.

Q So, chapter 8 of the Case Note Review Overview Report describes the areas of concern around – well, I can give you the headings, because that might be the easiest thing – data, managing infection control outbreaks, microbiology and IPC systems, clinical records of all sorts, adverse event reporting, morbidity and mortality reporting, line care, and other aspects of clinical care. So have you read that criticism?

A No.

Q No? Then the other groups we then move on to, I suppose, is into the two scenarios. So I'm going to leave the mechanics of ventilation to Dr Agrawal.

A Okay.

Q Do you see any connection between ventilation and the water and the environmental BSI that you're looking at in this chapter?

A No.

Q No. If we look at the water scenario, now when you wrote the report, did you know when the interventions took place?

A No.

Q Okay. So when do you

understand there to have been interventions?

A So-- All interventions or just water interventions?

Q Yes, water interventions. When do they start? When do they finish? (Inaudible – 14:30.55)----

A So it was early 2018. I believe it was March when they started concern about water-- when the concern about water was raised. In April, there was an intervention. Sort of April, May, June, there were a series of interventions from-- and I can't recall which matches up with which, but they first remove the taps, they then change out filters, they worry about drains, they dose drains, and then at some point in 2019 they're not satisfied, they want to look at the ward more and they move the patients.

Q So just to check, could it be that they don't remove the taps, they put filters on the taps?

A Filters on the taps, yeah.

Q They check the drains a lot, they dose the system two or three times, and then in September of '18, they move the patients.

A Okay, yes. That sounds accurate.

Q Then in December/November they turn on a chlorine dioxide dosing system which feeds the water for the whole hospital. Chlorine dioxide as a treatment agent. Is that consistent with what you think you read?

A Yes, yes.

Q Right. What be the mechanism by which such an intervention would be seen in the data?

A So what I would expect to see is – if that was the only intervention that was in play, right – that the cases would remain high and it would take a little bit of time, maybe two months, maybe three months, to start seeing those cases go down.

Q I appreciate you said if that was the only intervention.

A Yes.

Q The other point to mention is, imagine you have a water system which is being managed in a less than optimal manner for a number of years, and then in April 2018, the senior water managers at the hospital decide that there is contamination widespread through the water system. Now, whether other people later disagree, that's what is recorded in minutes. So the assumption there – and this is entirely an assumption – is that if there's something to be found in April, it was probably there for some months before. Now, if there was something wrong with the water – and that's an "if", I appreciate – in, say, the six months or the year before April '18, again, if that was the only cause, how

would you see that in the data?

A So, in the data you would again expect in the-- when it starts, and you hypothetically put that six months before April, you would probably see a small increase that goes up over time, and I would think that would lag the time that it started by a month, two months, maybe even three, depending on the mechanisms of how these micro-organisms are-- So now we're making a huge assumption, right, that the micro-organisms from the water are getting into the patients and causing bloodstream infections, so the mechanism for how that's happening, there will be a time lag for that.

Q Again, big assumption, set a scenario up: if we have a date, which we don't know when it is, but there's a date when the biofilm, I think is the word, gets to a point when it's causing it to slough off, and we've been told by some people that the risk is that it comes out the taps, it gets into eyes and mouths when washing the teeth and things, when you have showers, it gets into wounds, it gets onto things that are cleaned, including hands. Your position is that, within a few months of that date, whatever that date is, that is when you would see a slow increase in the data. Assume it's the only cause.

A Yeah, and you would probably

see certain micro-organisms come up first, right? So there are organisms like Pseudomonas that hang out, you know, in these biofilms that sort of get everywhere and create problems, and then you might see other types later on. I would expect to see a little bit of a pattern with the micro-organisms as well.

Q So I want you to answer on mechanisms, and I appreciate that you've got some caveats in there. So we then move to, I suppose, extreme positions, and what I want to do is ask for each of these – start at the bottom again – whether you can exclude them having no role at all based on the data you see, or being 100 per cent responsible. So it's excluding the extremes.

A Okay.

Q If it's in the middle, then obviously we'll do that in a different way.

A Okay.

Q So if we look at the lab systems contamination story, could that have no role at all or be 100 per cent responsible? Could it be at the extremes in that sense?

A So, if it's-- if it exists? Is that what we're saying?

Q Yes, if it exists.

A No, I don't think it fits either extreme.

Q Yes. So it either doesn't-- It can't be nothing to do with it----

A No.

Q -- and it can't be everything to do with it.

A That's correct.

Q Right. If we do the same thing for antibiotic resistance----

A Again----

Q -- I'm assuming that, given that not all antibiotics are affected, it can't be everything.

A Yeah, exactly.

Q Right. If we move to dysfunctional teams, I mean, the team can't be that dysfunctional, so it can't be everything----

A Of course not.

Q -- can it? If we think about line safety and the nursing and the single room, if we sort of fudge them together a bit, could that account for everything we see?

A (After a pause) So this is a very tricky question, and this is why I'm pausing. If we now make another assumption that it is only the things that we can attribute to being acquired in the hospital, it could, but remember that people can carry for years in their gut one of these micro-organisms, and it can translocate, yeah.

Q Okay, so----

A So it can----

Q I should rewrite my question about the 100 per cent as, "100 per cent of acquired in the hospital".

A 100 per cent of acquired in the hospital.

Q Could the line safety, nursing, room thing issue, could that cause 100 per cent of that part of what we see that is acquired in the hospital?

A I think it's unlikely, but it is the most likely of all our sets of-- So it is the-- it is the one thing from all our counterfactuals plus the water that I think is most likely to be responsible, only because this is what we know happens with bloodstream infections.

Q But you haven't studied the actual experience from the CLABSI report -

A That's right, and so I cannot say one way or another, and I don't think any of these would be 100 per cent of what we see.

Q So let's move to the sort of water scenario. Going to pass ventilation to one side. So again, holding your caveat that there's going to always be some that come from outside the hospital----

A Yeah.

Q -- of that unknown proportion that come from in the hospital, given what we see in both the environmental and non-environmental, could the water cause all of it?

A No.

Q No? And why is that?

A Well, for starters, we see our non-environmentals. Those don't come from the water.

Q Yes.

A And it's just, there's not a lot of biological plausibility with that, that water could cause this entire increase. So you have a small number of micro-organisms in the water. If----

Q So why do you say a small number of micro-organisms?

A Well, so-- So it is a-- Think of it this way, it's a large pool, right? If it's growing in the tap, for example, you have a lot of micro-organisms there, but when you turn on the tap, you have water flowing through that, right? So that's going to make that diffuse. You're diluting them, and we live in a world with micro-organisms in the taps all the time, everywhere, right?

Q What if it's in every tap?

A There are micro-organisms in every tap.

Q So you don't see a problem with that?

A Oh, I do see a problem with having----

Q No, I didn't mean problem in that sense. What I meant to say is, if you had a hospital, hypothetically, where biofilm was a problem in every tap, and maybe in quite a lot of dead legs, i.e.

(inaudible – 14:39.53) system.

A That is going to cause a problem.

Q You're saying that even then that couldn't cause the non-environmentals, because that's just not how it works biologically?

A No.

Q No? Right. So what I want to do now is play with dates, because if we go back to bundle 44, volume 5, and page-- Well, yes, let's start there----

A Sorry, my fault – that last question related to non-environmental----

Q Non-environmentals----

A Yes.

Q Non- environmentals cannot be caused by the water----

A Non-environmentals? Right, okay.

Q -- even if the water system had the problem in every single tap and all over the place.

A It is highly unlikely that the environmentals are being caused by that, because----

Q Why is that?

A Because predominantly you'll see Staph Aureus in here that is coming from the patient's skin or from someone's skin, because that's where it resides. It doesn't-- It doesn't survive in those environments, right?

Q The watery environments?

A Yeah. That's a low-nutrient environment with water that's not where it resides, and if you want lots of details about these micro-organisms, I'll refer you to Professor Hawkey. I'm sure you'll ask him lots of questions.

Q Yes. So what I want to do is I want to go back to page 50 on bundle 44, volume 5. Now, I think I need to have a final go at the dip----

A Yeah.

Q -- and so I'm going to enlist the help of Ms Cairns.

A Okay.

Q Visually, I'm going to take you to bundle 7, document 6, page 230. Now, just the top half of the page. So, this is a chart produced by HPS using their "environmental, including enteric group". Now, Ms Cairns' position is that this is the closest analogy to your group----

A Okay.

Q -- albeit it's not perfect. She's also given evidence – as has Ms Imrie, who have a part of the writing team – that you'll see how this chart plots a rate per 1,000 occupied bed days on the y-axis, time on the x, from mid-'13 through to late '19 when it's written. The blue line is an average, but that average is driven only by the data before the move to the new hospital.

A Okay.

Q And then what they've done is they've plotted an upper warning limit at two significant differences. At the point where it's-- The upper warning limit and the upper control limit at the second and third significance threshold-- I've forgotten the exact word, but it doesn't really matter for the purpose of my question. The reason I was interested in this chart is it contains data for the two years before the move. Now, it's just raw data, it's monthly, it's rate data, and visually you can see that it's bouncing around that short average from before the move.

A Mm-hmm.

Q So what I asked Ms Cairns to do was to do something that you don't like, which is to smish it all out and and get some rates, and so she did that, and it's in bundle 44, volume 6, at page 20. This chart is the same data as that table, and the chart we looked at, Figure 6, is the second big column from the right. Do you see how you've got four big columns at the top?

A Yeah, it's the environmental----

Q So they each have a figure in the chart, and then in each chart table, there is a "Cases" and there is a "Rate".

A Yes.

Q And it's smished, as you say it's averaged, across the period, so you lose all the detail----

A Yes.

Q -- and I asked her to, in a sense, mimic you, and she calculated a rate in Yorkhill prior to the move for those of 2.55, and then I asked her to count the rates each side of the decant, and she didn't like that very much because the third period is rather short, and so decided to calculate the year after the move when there were all those zeros----

A Mm-hmm.

Q -- and came up with 0.91. Now, if we go back to bundle 7, and that chart in Figure 6, I appreciate this is very short ranges of data, and it was done live at the time and I think NSS are quite keen for us to remember that. This is a dynamic piece of work done to support the IMT. But is there anything in here that helps us understand what's going on immediately before the move and that the start of the dip, this problem I keep putting back to you? There is actually a similar one in Dr Kennedy's work, which might be worth looking at, which is bundle 6. So, bundle 6----

A Before we jump over there----

Q Yes?

A -- can we look back at what Ms Cairns did most recently, that table?

Q Yes, of course. So that's bundle 44, volume 6, page 20.

A Because when I did look at this, this is one of the items you had me look at----

Q Yes?

A -- I was struck by the fact that it is following a similar pattern to what we're seeing with that dip, potentially.

Q Yes. Yes, but the dip-- well, I'll come back to the point I want to make when you've seen Kennedy.

A Okay.

Q So Kennedy was a public health doctor who was asked to produce a report by the medical director in 2018 and his report is in bundle 6, starting at page 95, but I want to show you the image, page 96. He called it, "Selected gram negatives." Now, the selected gram negatives were the ones that match the case definition for the IMT this was produced for. So it's a much tighter list than yours and it's possibly not entirely directly comparable. But his evidence was that this chart and its subsequent chart the following year showed three peaks. Two, which I'm not worried about today. I mean, one's in '17, one's in '18. But there's in July to September '13.

What I wanted to do if we go back to your long negative binomial chart at page 51 of bundle 44, volume 5, does this sort of contemporaneous sort of small scale, possibly quite quick and dirty data help you understand whether that dip is starting in early 14 because it's an artifact of the GAM or because that's when it's actually starting or is it, actually, just

dealing with the fact that there's a change of system into a new water system on the move, and actually although I'm very bad for-- I keep pushing this as a possibility to you that actually the change is the move, that's the point of change. Does that help you do that or am I just putting words in your mouth?

A So if I understand you correctly----

Q Yes.

A -- and please clarify if I don't, are you saying that there is a water system at Yorkhill----

Q There is.

A -- and it's probably old because that's an old hospital and it could have all sorts of-- nobody tested it so we don't know what's in that water.

Q No-one really knows, we haven't had evidence of that.

A And then you're moving to a new hospital and you're making an assumption that the water is clean when they get there, and it's not been used and----

Q No, I'm making an assumption that it's not necessarily the same. I'm just assuming----

A The water is just different?

Q -- that it's different. There's a good-- I'll rephrase that. Is it reasonable to say there is a chance that there is going to be a change just because there's a change? Not better or worse, just different.

A That's probably beyond our ability to hypothesise against the statistics we're seeing. I mean, we could say-- so if we're throwing out, like, counterfactuals.

Q Yes. Because there's no counterfactuals apart from the single rooms, so that is important.

A Yes. Yes, I----

Q They obviously have a change point here.

A I mean, if we start talking about counterfactuals, why there would be a change point there, maybe it's just moving to a new environment. Maybe it's single rooms. Maybe there's extra attentive care being taken to the patients because of the move. You know, I'm just coming up with ideas now about----

Q Yes, but if we have evidence that you weren't shown about-- well, you've now seen some of it about how the CLABSI issue emerges. We should probably talk about the CLABSI evidence. So you read Jennifer Rodgers' statement?

A I did.

Q And I think you might have looked at a chart that she produced with-- where's the document list? But you read Jennifer Rodgers' evidence?

A Yes.

Q Now, what I'm wondering here is if we start with a non-environmental, I mean, that's the next page please, page 51. Do we see any reason to think that the story that is described by Ms Rodgers might be contributing to this non-environmental curve and the way it's behaving?

A Yes, I think that her story could suggest both curves in that it-- so if they are actively monitoring, and I'm not sure how actively they were monitoring because she doesn't start there, right? She says someone noticed in early 2016 that there seems to be an increase in bloodstream infections, I think the CLABSI infections, and then they start monitoring it quite intensively. Usually, when you're not monitoring things and you sort of get a clinical feeling for them, it's after they've started quite a while, and so you could say, "Well, you know," and if we want to play the game of percentages, a lot of that could be due to them having a problem with their line handling.

Q So I suppose that brings us to the crux, the crunch as it were. I know you didn't read the Case Note Review and we can't rely on this number because we can't see the workings, but in many ways it might give us an upper limit. The Case Note Review says, if I can remember correctly, that 30 per cent of the cases they looked at, all of which were in the environmental side of your chart, were more likely than not to have an environmental connection. Are you aware of that?

A I'm aware of that. I'm not aware of the evidence they put (inaudible).

Q I appreciate that, but the reason I wanted to pick on it is just to pick as an upper limit, and then you'll be aware that it's the position of Greater Glasgow and Clyde Health Board that only two of all the cases can be associated with the environment. I suppose that's a lower limit.

A Yes.

Q And you've just given evidence that, of the non-environmental cases, quite a high proportion might be related to line safety.

A They might.

Q And about 10 minutes ago you gave evidence that we should remember that in both groups there's always going to be a chunk that's brought in from outside.

A Yes.

Q And so what I'm wondering is, I appreciate you're not involved in-- you explained you're not involved in what can only be described as the disagreement between Professor Hawkey and Dr Agrawal on the case notes review, but there does seem to be a disagreement.

But could I put this to you, that what we see here is consistent with, not just in page 51 but in page 50, it's consistent with the CLABSI having an impact on non-environmental, the water having an impact to some degree, possibly in that lower third range on the patient for the environmental. I mean, not nothing but not a hundred per cent, and then what gets brought in from outside also has an impact. Is that broadly a happy position to be in?

A Yes. Yeah, I would agree with that. So remember that environments, regardless of whether or not we see them as, you know, extra contamination in water or just the normal amount that you get from water coming out of, you know, the mains that you're given, it's part of the environment, right? That's where people would pick up environmental organisms. It's just a matter of, is it the hospital? Is it at home? Is it-- you know? So there is no question that the environment plays a role.

Q Yes. I mean, I wonder if we can take you back to the HAD report and to page 118 of 44, volume 1, where I think you reached conclusions on what your chapters have showed. So yes, page 118, please. So figure 22 is at the top, and then we have a heading, I appreciate you didn't write the headings:

A "Evidence of environmental contamination, particularly the water supply, based on bacteraemia data from haemato-oncology patients in GGC. In summary, the historical data from the GGC haemato-oncology services provides a background into prior incidence and cases of bact6eraemia attributable to microorganisms that could be environmentally relevant. This allows us to evaluate cases and incidence at QEUH relative to the same or equivalent services for haemato-oncology patients within the local Glasgow region."

Q Now, was that actually what you were doing in 7.2 for the children?

A Yeah, and I think that, more specifically, we are talking about could the environmental data explain everything? Does it show that there is an increase and is explained by the water?

Q Is that the question you were asked? Because if we go back to Dr Agrawal's instruction letter, which is on the appendix page 242 of the same bundle. The question 4, 5, which is the one you were asked, 5A is, "Is the water testing results consistent with there being widespread contamination?" Can you explain how you tested that without knowing when the widespread contamination was supposed to be there?

A Yeah, and I don't think that was tested. I think we really focused on, is there an increased level of infection

consistent with widespread contamination?

Q Without knowing when it is?

A Yes, well, we were given the impression that it was from the beginning.

Q Right, because if we go back to page 118, there's a series of bullet points, and I really want to understand whether you currently accept them. So the first bullet point might seem to be a truism but:

> "Haemato-oncology patients across services have an expected, yet very high risk of bacteraemia and cases are common in this patient population."

Would you agree with that?

A Yeah.

Q

> "Among haemato-oncology patients, including comparison of both historical and current services, both incidence and cases of bacteraemia attributed to environmentally relevant microorganisms individually taken together are statistically similar between the QEUH and other GGC hospitals. "

Would you agree with that?

A And I believe we're looking at adults right now, right? Because I've got----

Q No, this is adults you think? Okay.

A This is----

Q Right. So would you agree with that?

A Yes.

Q Right, and is that in a sense because, apart from that spike in 2018, there was no difference?

A So for the adults, there isn't much difference. We didn't go through that in detail. And for the paediatric, I would definitely say I would revise that bullet to say----

Q Well, it's over the next page. We might want to look at it.

A Okay.

Q So go to the next page.

A No, that's the first one's adults.

Q Yes. The second one also, we've discussed how they weren't exposed to the water. So the second one is-- and does that help us understand, if it's true that for:

> "The most vulnerable adult patients in this service, bone marrow transplant patients, alone there is a continued decrease in incidence overtime following transfer of these services from the Beatson Hospital in the North Sector to QEUH."

A Well, understanding that they received the intervention water, if that's what we want to call it, right?

Q But they never received the widespread contamination?

A Yeah, they never received the contaminated water, if that's what we're calling it.

Q So what I'm wondering is whether that bullet point really helps us?

A What it helps us in saying is that certainly what we expect, so if we had the South patients to compare, what we do expect is a decrease in this population if water isn't contributing, right? Because this is our patient population that moved from North to QEUH and did not receive, as you've stated it, contaminated water.

Q Yes. So there wouldn't be a change or it would get better? It wouldn't get worse.

A Yes, it certainly wouldn't get worse. It would get-- I would hope it would get better because that's actually what we're seeing.

Q Right, and then the second bullet point:

> "Among paediatric haemato-oncology patients, we see an ~2 fold decrease in incidence and cases of bacteraemia attributed to environmentally relevant microorganisms following transfer of services from Yorkhill to QEUH. "

A And this is the one I would edit.

Q So how would you edit it?

A I would say from the beginning, right before transfer to the very end of our data set----

Q Yes?

A -- we see that twofold difference, right, and that's shown in our line of decrease, but in the middle there was an definite increase in cases for an acute period of time.

Q And would you give that increase a period of time it happened in?

A Yes, I would.

Q What would that period be?

A Can we go back to my graph?

Q Well, we can jump one page back to 118 or we can go forward to the GAM one instead.

A Can we have the GAM one instead?

Q Yes, so that's 44, volume 5, page 50, I think.

A Page above.

Q No. Page 50, yes.

A A page above.

Q There we are.

A Yes, I would say the increase was from roughly, and I'd have to look at my analyses, that looks like about March 2016 through to-- I'm going to guess that's August 2020, but I have the dates in there. I'm just-- March to-- No, March to----

Q I think it might be----

A Oh, I only called the downward trend. I'm sorry. Yeah. So there is a downward trend from January 2018, so there's----

Q But the increase is the second quarter of '16.

A Yeah. There's an increase-- If we're going to do this by increase and decrease, there's an increase from March 2015 to January 2018----

Q 2015 or '16?

A '16, sorry, and then there's a decrease from 2018 to a fairly low level in what we would like to see, right, as an overall decrease.

Q Would it come to any surprise to you that Mr Mookerjee's position was we saw an increase from the second quarter of 2016, reducing again in 2019?

A No, it's a little late, the 2019, but it's not a surprise.

Q Now, I haven't discussed clustering with you because I was a little confused.

A No.

Q So you carried out a clustering exercise in the report. Was that your work, the original one?

A So, I did the work. It was suggested to me-- I was never a fan of doing a qualitative sort of thing but I was a willing party in working on that.

Q So, what's wrong with doing the qualitative clustering exercise in the HAD Report?

A So, the problem is it's-- There's nothing wrong, per se, with qualitative items, right? You see that a lot in, you know, root cause analysis and other things, right? You interview people, you find out what they think is going on. This is all very qualitative. There's nothing wrong with doing qualitative things but I'm not sure how much it helps, because what we're saying is, "Okay, this type of organism we see in this month, we see it in different wards, we see it in the next month," and then we're sort of setting a case definition for how close together in time they can be and what we call that.

And so whilst that might be interesting, there's no-- So, let's take something like Klebsiella, which is one of our very common organisms that we've seen. There's no information for us that those are clonal or truly clustered, right? They are occurring together in time, as we're calling them clusters, and it's probably-- you notice in the rebuttal, I move away from the concept of clusters. I don't actually really like that. They're runs, so things occurring closely together. It doesn't mean that they're related at all.

Q The reason I ask that, I'm intrigued with the idea that you had some concerns and then you did this entirely statistics-based exercise, which is literally

the numbers driving it, if I understand correctly.

A So, the GAM, you mean? Oh, you mean this next one-- the next one?

Q The clustering exercise, yes. That's basically just the numbers drive the process.

A Yes.

Q Yes, so sticking with the report, the report contains a statement that we shouldn't-- Let me find the exact words. At the end of the introduction of the report that you signed on page 17 of bundle 44, volume 1, it says:

> "A significant reliance appears to have been placed on the Queen Elizabeth University Hospital and Royal Hospital for Children [placed on the Case Notes Review] … in relation to whether the built environment at the QEUH and RHC posed an increased risk of infection. However, we do not consider that the CNR is of assistance in determining this for the reasons set out below."

You didn't read the CNR Overview Report. How could you sign that?

A Yes, that was-- Yes, that was from Professor Hawkey.

Q Because what's worrying me is what to do with your latest statistical methodology. Because your report was produced, I appreciate that it went through a procedural journey to get into the Inquiry. We instructed you in February/March. We sent your report to the Case Note Review authors. They wrote a rebuttal, which they hadn't done before, and then you've done a completely new exercise, which they haven't seen. What am I supposed to do with your completely new exercise? Did you read the Case Note Review before you did your completely new exercise?

A No.

Q Can you give me a reason why I shouldn't just put it to one side as an exercise carried out entirely in the dark without looking at the people you're criticising? Or are you not criticising the CNR?

A So, I am personally not involved in criticising that. I would ask you to direct those questions to Professor Hawkey. I think he has a very clear set of arguments that he's interested in. I would say the new analyses that I produced were largely in response to valuable commentary from Ms Cairns.

Q Right. Now, if we go to your new analysis, which is in the Overview Report-- I mustn't forget where I am. I'm going to have to put a marker in here before I lose it. Volume 5, the rebuttal, the response document. Your new attempt is in bundle 44. It's in this

document. Allow me a moment just to get my hand on it. Page 65, please. No, it's not. Sorry, it starts at 54. So, you carry out a clustering analysis by a new method.

A That is correct.

Q I just wondered why you decided to carry out the new method.

A So, there is still the concern about what is happening, right, is-- What are we seeing here, and how do we sort of help ourselves understand? And so that is, from an epidemiological and statistical point of view, you think about different ways to model the data.

Q Would it be fair to say the conclusion of this exercise is, you don't actually see clusters, really?

A Runs, yes, runs.

Q Runs.

A You do in a few of the organisms, but we don't see a lot of runs.

Q So, what I want to do is to move to a topic that has always confused me, and I wonder if you'll indulge me by just me setting out a scenario. Let's imagine that the reality is, and purely imaginary, that some proportion of the infections that were observed for both the environmental and non-environmental in the paediatric patients were brought in from outside.

A Okay.

Q And when I say "some," in all

cases, I mean a non-negligible amount, and that some in the environmental were caused by widespread contamination of the water system, and that some were caused by gut breakthrough or colonised patients from the environmental group. Amongst the non-environmental group, some were brought in from outside, some were caused by line care problems, and some were colonised patients as well. So, in effect, we have seven different causes spread over two types of bacteria.

Now, if that was the case, and I realise the sum is doing an awful lot of work in that paragraph, would this sort of method of looking for runs actually find anything?

A So, if that description of those seven items was completely chaotic, in other words, there's no pattern----

Q No, I think there has to be a pattern. I can't say some water and some Klebsiella unless there is a pattern because they have patterns, in that we can see patterns if they're true.

A So, where we would likely see patterns is, for example, if-- and this is what you get with Pseudomonas, one type of organism growing in the town, or some contaminant in the environment and poor line handling, and they keep touching whatever that contaminant is, that's where you would see something in those runs, because you would get

repeated cases of the same organism, but you would have also seen that in the sequencing if they were able to do a lot of whole genome sequencing, right, because you'd see clonal organisms.

Q Yes, but if we just look at the cluster exercise, what's concerning me, I suppose – it concerns me and pretty much everybody who's given evidence from an expertise point of view – is that often it is said I can show, yes or no, whether something is the cause. Of course, is there really ever going to be one cause?

A It is highly unlikely.

Q Yes, and so if we go back to your bullet points on page 119 of 44, volume 1, what I'm just wondering about – the previous statement about clusters, your work on clusters, and the Case Note Review work on clusters, which they did through a qualitative analysis – is, in a sense, is it entirely possible they're just all right? Because they're all coming at it from a different direction, doing different things. So Professor Hawkey's approach is qualitative, but with very little material, just dates and wards. The CNR is qualitative, but vast amounts of data, which we can't see. Yours is statistical, just dates, really, nothing else, maybe. Actually, they all come at it and they just slightly miss each other. They're not actually contradicting each other.

A Yeah, I guess this is what I would liken to the approach that I keep mentioning. If you model the same data with different statistics and you get the same result, that tells you something. If you get very different results, you're left with still a lot of questions and not a lot of answers. And I guess that's where we would be left. I don't know that these qualitative approaches are as comparable.

Q Right. If we look at the second last bullet point, overall patterns of:

> "BSIs attributable to environmentally relevant organisms largely follow the same pattern as BSIs attributable to organisms that do not persist in the environment."

Do they follow the same pattern in the children?

A No, I would change that to a similar pattern where environmental starts-- non-environmental starts first, and I would say, "Cannot be fully attributed to the environmental," because there-- I don't think-- and I don't want to speak for my colleagues, but I don't think any of us believe that any part of the environment cannot be attributed to some cases. What our understanding was, is that the entire increase was being attributed to-- and maybe that was a misrepresentation

given to us, to water, which is a very big stretch.

Q I'm going to challenge you on the similar. Can we go back to bundle 44, volume 5, page 50, please? Now, I'm going to look at the two charts. This chart has a point when the data is above the blue line and a point when it's below the blue line. The next chart, page 51, at the point before the increase, it isn't really below the line in a sort of true sense. So is that a difference?

A Yeah, that's different.

Q And do we need to think about why that might be? Or do we just put that out of experience and think it's not important?

A I think I very much appreciate that you want to think about that. I think it's interesting, and I think largely I ignored it because the focus, of course, was on increased cases.

Q Because it just occurs to me that if we look at the non-environmental one, and we start thinking about CLABSI -

A Mm-hmm.

Q I suppose line care generally can get better, but it generally gets better after it's got worse, because you have to work at it to get it better, so I'm wondering whether a peak happening from the linear without a dip perhaps isn't that surprising if it's line care, because things just get worse for a reason we can't necessarily find or a time-- and then they (inaudible 15:15.29) it and it goes down again. In fact, it drops below because they're really efficient. Whereas if we go back to page 50, if it's the environment and you put yourself into a nice clean environment for a short period of time, you get better. Then, of course, if you don't manage that nice clean environment, it gets worse. And so are those shapes perhaps telling us a story about out the direction of the change pressure in the two different types of microorganisms?

A So, they could be. I want to note that it starts decreasing in 2014, which is before the move.

Q Which is confusing on that theory.

A So I would say certainly that there might be something about the different microorganisms and the pressure on the microorganisms, but if we're-- So again, I'm perplexed that water would be necessarily proposed unless there's something that happened at Yorkhill that we're unaware of.

Q Yes, so we don't have any evidence of that. I think this is a reasonable summary. They move into the new hospital. There is lots of noise about ventilation in the new hospital, but there isn't really any discussion about water in the infection control team for at

least a year, but there is lots of discussion about ventilation, and that's why there's a contrast, I would say. Although we have the DMA Canyon report, which I think you've now read, full of anxiety about the water system, I'm wondering whether the existence of a drop below before the increase might well be consistent with, for a period of time, the water being cleaner than it was at Yorkhill, and that might suggest that the change of the buildings is the change of the drop.

A So, what I'm going to say is it could be a reasonable hypothesis. What are your counterfactuals?

Q So counterfactuals are you move into a new building, it's absolutely spiffingly clean, everyone's very happy and they don't have any issues of bad infection management and control in the single rooms, but that counterfactual is inconsistent with the environmental growth because the environmentals were going up at that point. So if the experience in the single rooms is good in the new hospital, as I suppose you'd hope it would be, you wouldn't see the increase in the environmental at the same time as a decrease in the environmental. It's that contrast that I was trying to explore.

A Yeah, I see what you're saying. So you mean if it's line care

there?

Q Then both of them would be going up.

A I guess I would say that is true, although I might say that a counter-argument would be if it is line care and you go into a spotlessly clean hospital-- Let's assume it is.

Q I mean, I don't think it was, but we'll assume it is.

A Yeah, let's just assume it was. Then they're not picking up anything from the environment on their gloves or wherever to contaminate the lines----

Q Yes, so you would only get non-environmental increasing.

A Yes.

Q Right. So that's a different way of seeing it. Okay.

A Yeah, that's a different way of seeing it. So that that could be one option. I would be very cautious to describe everything that we see as that.

Q No. This is my concern, that, I suppose, as we learn in the Inquiry about stuff, the absolutes seem difficult to understand. But you seem to have rejected the absolutes, so we can-- It's not all the water, it's not all the line care, it's not all the management, not all these things. Let's go back to the bullet points on page 119 of 44, volume 1. I just wonder about the last bullet point.

> "There have been large reductions in incident BSI overall and by microorganisms environmental attribution at QEUH overtime(sic), suggesting overall increasingly effective prevention strategies. "

I appreciate that was shown by the linear regression from figure 22, the lines not on the chart, but is it a good idea for a panel of experts to say that when they know that the case notes review has as a principle conclusion that there were problems with infection prevention and control in the hospital at that time? And that's not their clustering. That's their review of the evidence.

A Yeah, and I think that it could have been better pointed out that even if we look at the new modelling of the latest data-- I can't say currently because we don't have data from current. The latest data we had showed, regardless of whether or not there was a problem in the middle, which I completely agree with, that things are better now than they were before. For example, if we take Yorkhill, they do have lower rates now than they did when they left Yorkhill.

Q I suppose the final counterfactual point before we move on to Aspergillus is this. There seems to be a view, and I wonder if you can comment on it, that the two facts that Ward 2A and Ward 2B were fully refitted at a vast expense with a ventilation system, new water pipes, considerable effort made to make sure the water was clean, including replacing taps and a lot more work than perhaps even was expected, and now the rates for environmental and non-environmental are low, and in fact for non-environmental lower than the trend-- Whilst counterfactuals are important, the fact that you had to replace the entire ward tends to suggest the environment might have been part of the cause.

A It does suggest that. I'm going to throw out a counterfactual that I do think is really important here. Once it gets a ton of media and press attention, then people are gonna be pretty paranoid about how they look after lines and what's going on and prevention in a way they might not have been focused on before as well. So I think----

Q You think in the sense we've got to remember that things will always get better anyway if you apply attention to them.

A Yes. Yes, they will, and that's not to say that I don't think that refitting the water could have contributed to that, but I just want to highlight that when people are under a microscope, they behave in the way that they think is expected of them.

Q Right. What I'd like to do now,

my Lord, is move on to Aspergillus, which is my final major topic, and then we'll have our 10-minute break for Rule 9s, once I've done that. I think we might run a little bit over. I have warned my colleagues, but I think I should mention that.

THE CHAIR: So what do you suggest?

MR MACKINTOSH: I think I'll be going now until about five to, so we'll probably be looking at a half past four finish, I expect.

THE CHAIR: Right. So, sorry, I may have picked you up incorrectly. Are you looking for a break now?

MR MACKINTOSH: No, my Lord, I'm just briefly warning my colleagues because I discussed it briefly before lunch and I wanted to make sure they saw where I was.

THE CHAIR: Dr Drumright, you're quite comfortable?

A I am. Thank you.

MR MACKINTOSH: So, what I want to do now is to move on to Aspergillus, and I suspect what I have to do for chapter 8 is talk to Dr Agrawal, really, because he will do the work.

A Yes, but I am happy to answer the questions about the new modelling.

Q Yes. So what I want to do is briefly look at chapter 8 just to get context, then look at the new modelling, and just for completeness look at the new modelling that we asked you to do and then ask you where it takes us. So if we go to 44, volume 1 and we go to chapter 8, which is on page 123. Now, obviously you didn't do the analysis of the imaging. You're nodding again.

A Yes, I did not do the analysis of the imaging. I did not call any of the cases. That is all Dr Agrawal's speciality.

Q So he gave you counts.

A Yes.

Q So in fact, if we go to page 125, what we see in figure 23 on page 125 is the counts for the adults by year.

A Yes. I would have to double check that that is the final set of counts I had when I analyzed the data----

Q I think you made one change. I can't remember which year it is, but it's the counts at the point he wrote the chapter.

A Yes.

Q Yes. And we see Adult South is blue, and apart from 2018-- Well, in fact, before 2017, we have years that are above 1, but otherwise it's 1 for every year through the data field. And then Adult North is in orange and Adult BMT is in green, and I suppose we should also notice the bars he's put at the top to show when they were in the hospital, the green ones.

A When they were in QEUH,

yes.

Q Yes. And our understanding is that from 13 June 2018 they were in HEPA filtration positive pressure to the rest of the hospital, six air changes an hour, in ward 4B. But when they were in there briefly, in June/July 2015, they were not in such a bespoke environment. Now, he didn't calculate a rate. Is there any particular reason that you know about, or should I just ask him why he didn't calculate a rate?

A I think you should ask him. I think my understanding is that the numbers were very low, so.

Q If we go on to page 128, he's plotted the paediatric haemato-oncology rates, similar style, only one column, and he's plotted the locations within the children's hospital above. So we have a bar between June '15 and September '18 showing they're in 2A and 2B, and then we have the bone marrow treatment patients in 4B, and the in-patients, and indeed the out-patients as well, in 6A until '22, and then everyone's back in 2A and 2B. And I'm assuming the same issue is the reason he didn't plot the plot the rates (inaudible).

A That is my understanding.

Q Now, I asked Dr Mumford and Mr Mookerjee to calculate rates and they did that. I'm not proposing to take you to them, because you eventually did your own rates.

A I did.

Q And although I think NSS have some criticisms of Dr Agrawal's methodology to construct these rates, he has to get some credit for constructing some rates. No one else has done it. So we'll use the rates that you constructed in the calculations in the response document. It's bundle 44, volume 5, and we go to page 66. Now, what I wanted to do with this was to firstly check in with you about methodology of constructing these rates.

A Yep.

Q Because I suppose we can end up down the same rabbit hole as we were before with the BSI. Has he effectively used the same occupied bed data?

A That is correct. That is the occupied bed data I had.

Q So if there is a mismatch, it's the same mismatch?

A Yes.

Yes.

Q Yes. But with that in mind, we should probably go and look at a particular chart that he produced, because it's an annual chart, which is on page 69. And I suppose the reason I wanted to go to that, although it's a different style, it's roughly the same data as page 124, which we looked at before.

It's the three different parts with annual rates.

A Yes.

Q Now, from onwards in both exercises, you moved to monthly rates. Why did you move to monthly rates?

A So again, this is a this is a preference for data analysis. When you remove data, so when you go from month to year, you're removing information that is valuable and you're removing power. So the reason that it's graphed on yearly rates is because that was our understanding of what was asked for us as well as the monthly, which is my preference.

Q But the original count was just a count, so it didn't matter about power. If we move on, for example, to look at what could be described as a confusing graphic on page 71, the very number of zero months.

A Are you looking at the table?

Q I'm looking at the chart.

A The chart, yes.

Q I mean, the table works as well, actually, to the point.

A Yes.

Q There are lots of zero months. Does this actually get unhelpful at some point?

A Not necessarily, no, because if you ignore those zero months, you're losing information again. So remember we talked about the negative-- the negative binomial GLMs?

Q Yes.

A We can use those with this and, for example, what we actually see in here is not that any one month – with the exception of that that peak in the middle – is really gaining a whole lot of-- of cases in a month, right? We see one, two, or zero cases per month, always.

Q Yes.

A It never goes above two, but what we see later on, and we see our curve is-- when we look at the whole set of data, we do have a significant linear increasing trend. Very slowly increasing----

Q Yes.

A -- but increasing. Our zero months are just getting fewer.

Q Well, if we go on to page 73, we see Aspergillus. This is you going back to your negative binomial----

A Yes.

Q -- GAM plot.

A Yes.

Q So, I mean, at one level, this is difficult to understand. The dots are----

A Mostly zeros----

Q I mean, the human eye sees the dots that are on the line, but of course you've got to remember there are lots below the line. So what's this chart showing us?

A That chart is showing a-- And this is-- I just want to make sure. This is-- So I want to make sure that everyone's aware: this chart right here is what you asked for, where we cut up the data.

Q No, no, this isn't the----

A Oh, is this the original?

Q Yes.

A This chart is showing that there's a linear trend. Your smooth line is completely over the top, that's why it looks purple, and I believe-- And I want to double check-- Oh, no. The P value-- Let me just double check----

Q It says at the bottom of the page there's no significant change over time.

A No significant change over time, that's showing.

Q So that's the adults at the Queen Elizabeth.

A Yes, that's the adults.

Q No significant change over time.

A Thank you. Adults have no change. It is a flat line. So, it is saying that adults basically have Aspergillus largely at the same rate over this whole time period.

Q Yes, and if we look on the next page for the other hospitals that you've got data for----

A It's also flat.

Q -- it's also a flat line. Right.

A And they're roughly the same.

Q So, in a sense, that doesn't really tell us anything, other than there's not really a trend.

A Well, what it tells us is there's no difference between these groups.

Q Yes, yes. Even though some of them have ventilation that's better than others, it's not making any difference?

A It's not making any difference, yes.

Q Then we go to the paediatrics. That's the next page, 75. We have on page 75 Figure 5C.1, the monthly counts, and then on page 76 we have the annual ones.

A Yes.

Q Next chart, next page, and that's at 5C.2. So your comments, I'm assuming, about the zeros and the binomials apply the same here?

A Yes.

Q Right. Now, over the page, you plotted the paediatric Aspergillus on one chart.

A Yes.

Q What does this GAM negative binomial line show us?

A And I apologise, this is when we went to the adults. I thought we were going to talk about paediatrics. What this is showing is a very, very gradual increase over time that was statistically significant.

Q Is it possible to see in this chart whether there is either any change point at all or when that change point is?

A So this chart right here suggests no change point, and there's a subsequent analysis in the questions you sent to Mr Mookerjee and I where we don't see a significant change point, but I did, as you are aware, some other modelling to sort of impose a change point. So that-- that's in those data. I don't know if you want to go there now.

Q Well, what I wanted to do was just to walk us through these slowly----

A Okay.

Q -- because I suspect we're going to get nowhere in the end, but I want to just see where we get to.

A Sure. No problem.

Q Of these two, the linear and linear smooth, are either of them more significant than the other?

A No, I believe the linear-- the linear smooth isn't significant because it's matching the linear. So there's----

Q What about the----

A That little----

Q -- feature on the right-hand end where it starts dipping down?

A That little dip down on the end, yeah, it's not going to call that.

Q Okay. If we go on to the-- I think this may be, page 80, where you imposed a change point. Is that where you----

A No, this-- this on page 80, I took-- So this is based on annual data now. I took----

Q Right.

A Because that's the only data I had from Mr Mookerjee.

Q Yes.

A I took his data and I plotted it, and I took our data and plotted it, and that's what you can see. The HAD is the red and the MM is the blue, and it was just to demonstrate that, even when we have this wildly different number of cases at the very end, most of our case numbers and rates are-- are about the same.

Q Yes.

A But at the very end they're wildly different and we're still seeing the same-- the exact same trend.

Q Same trend? So you wouldn't think there was a-- In the PDF-- Well, let's go and see what Mr Mookerjee did----

A Yes.

Q -- because I think you need to comment on that so that then ultimately he can comment on you. So we go back in the same document. Well actually, no, let's not do that. Let's go look at what I asked you to do first. So that's volume 7. I think I've obviously been told quite carefully by you what you think about

chopping up data points, but we should probably look at page 55. Let's not do that, let's look at page 53. So I asked you to look at the Aspergillus at Yorkhill by itself. What's happening here?

A So what you see is your blue line is always your linear and your red is your smoothed, and in this case we do see this increase in the smooth line, but that line is not statistically significant, so----

Q So we should put it to one side?

A So we put it to one side and what we see-- And if you can scroll down, it's underneath. I believe----

Q I think it might not be----

A -- that trend was not significant----

Q Over the next page, actually, 54.

A Yeah. I think----

Q The trend there isn't significant?

A Is not significant, yeah.

Q So even the blue line isn't significant?

A No.

Q No, right.

A So what that says is there's----

Q No trend.

A Yeah.

Q But in the context of a longer overall gradual trend.

A So no trend in that chopped----

Q So that bit of Yorkhill, but we've got to remember that the whole long sequence has a small----

A Does have a small upward trend.

Q Right. Then we go to the similar short chop for Yorkhill on page 50-- Not Yorkhill, Queen Elizabeth, on page 55. Again, what does your GAM exercise tell us here?

A So again, the smoothed line is not significant, so we should ignore it and go with the linear trend, and again that's not significant.

Q So the only thing that's been significant is the long trend over the whole period?

A Yeah, let me just-- Yes, over the long period.

Q And in fact, helpfully provided that on the next page, at page 56.

A Yeah, yeah. There it is.

Q So let's look at Mr Mookerjee briefly. He decided to do a linear trend line for Aspergillus.

A A linear regression trend line, yes.

Q And he plotted that on page 42, and I understand you have a little bit of a concern about his confidence intervals.

A Yes.

Q So what's your concern about

Figure 1 on page 42?

A So-- How interesting. His confidence intervals are showing up a little bit-- His confidence intervals are awfully tight.

Q You mean they're very close to each other?

A Yes. So when----

Q I wonder if we could zoom into the bottom half of the page, just to make it a little bit easier for us. (After a pause) Yes.

A When you have that variable of data, your confidence intervals get pretty wide. When you chop up your data and you reduce your power, your confidence intervals get pretty wide. I tried to model what I saw here. I could not replicate these confidence intervals.

Q So you think that----

A I'm just not----

Q You would disagree with Mr Mookerjee and say that the idea that there's a step change at the change between the buildings doesn't have any statistical significance?

A Well, that and the confidence intervals overlap, right? Do you see right there the yellow and the blue?

Q I do.

A So I think it's largely-- The modelling I would also argue is not correct, because this is using not a GLM but a linear regression----

Q And so it won't deal well with the zeros?

A That's right.

Q Right. Let's take that off the screen to go back to the HAD report itself.

A And just to highlight again: if confidence intervals overlap, then that is a possible point. So they possibly could have the exact same point estimate.

Q Exactly. Yes, I understand.

A Yeah.

Q Can we go back to bundle 44 volume 1, page 128? I suppose where I'm sitting at is I'm stuck and I need your help.

A Okay.

Q So at the end of this page, Dr Agrawal – and I'll ask him about it too – says, bottom of the page, 128:

> "Table 19 shows the breakdown of cases by location throughout the period the paediatric service was based on Wards 2A/B, RHC, and during the time of relocation to... 4B and 6A. The annual total number of cases for the paediatric services varied from 1-7 (2015-2022) compared to 0-5 cases per year prior to 2015. No trends are seen over time... by location when the service was split between... 4B and 6A."

Of course, you went ahead and, using what bed day data you could find, you concluded that there's no statistical significance between the two periods. In fact, there's a gentle trend upwards, but that's all. Obviously this is a public inquiry, and we're not necessarily answering the question of, "Is there a statistical significance in this data, or should some scientists reach these conclusions?" We've got to work out what to do about ventilation that doesn't meet guidance. What role should this chart – and I recognise it's only numbers, not incidence, and the incidence is not statistically significant – play in our deliberations of, "Does it matter whether you follow guidance or not?"

A I want to be a little bit careful about this.

Q Absolutely.

A I'm going to back up and say I might not answer that question directly. There's lots of reasons to follow guidance----

Q Yes, of course.

A -- and it stretches outside of this Inquiry, I would suspect. I could be wrong. The question of, "Is there a problem being caused by ventilation that doesn't mean guidance?" I think is-- is more of what this is pointing toward, and it is suggesting maybe not, although again we should probably understand--

and I don't know if you want to talk about this when we talk about my modelling of where a change does occur in the data----

Q Well, I was going to come to that in a moment, but I just felt that the question could actually be asked both ways around. So you could ask the question-- or maybe we come back to it after looking at the change, is, "Does this piece of information help us to determine that you don't need to worry about the ventilation being below guidance?" or, "Does this chart help us determine that do need to worry?" So it's sort of opposite way round. Does it help either way in proving that we need to worry about ventilation system?

A I'm going to not directly answer that question again and give you a different plausible answer, which-- this information plus a lot of literature out there which I know a portion of which shows that ventilation at the rate we think doesn't have the impact we think and doesn't do the things we think. I think what it suggests is, there is a lot of need to explore and understand what is going on with ventilation and what should be proper ventilation in healthcare settings. I'm not sure that current guidance meets that. I'm not sure that this is the right way. What I'm suggesting is that, we know very little about ventilation, and

that's-- to keep going on with that is probably not a great idea.

Q Okay. Well let's go back to your calculations on change.

A Yeah.

Q So 44, volume 7, page 65. So it's sort of section 3 of your report.

A And this is our response to you now, isn't it?

Q Yes, it's in the-- well, this is your document.

A Yeah. So I mean, I can summarise it if we don't move along, but there is a table.

Q Yes, if we go on to page 68.

A Yes. So in this table, what I did was I was curious if we look at my model, okay, so we put aside Sid's model and we look at a linear change model which is the one we saw before, the GAM model that it's increasing slowly----

Q Yes.

A -- and we tell the model it has to impose a break somewhere.

Q So it's a bit like when I gave you the questions.

A Yes, but I told the model to do it right based on the data.

Q Right. Yeah.

A It picks the point in 2019, not at the move.

Q Right.

A So----

Q And how do we see that from the table?

A Okay, so from the table, the table's not showing you that. What the table is showing you is comparability in models. So I throw in a model in there that I know is no good, which is the constant, it assumes nothing's happening, and that model has an AICC weight that is quite low and that's showing that's not comparable. But if you look----

Q So you're comparing a particular model to the data?

A We're comparing the models to each other.

Q To each other, right, yeah.

A Yeah, and so that model is completely out. But if we look at the linear change model, which means there's a constant linear change over time----

Q Yes.

A -- versus the step change model that picked a date in 2019----

Q Yes.

-- those models, you can see that the step change model is ever so slightly better but at .11 in the difference, right, that's not considered significant. You need a difference of two or more to say a model is better.

Right.

A So what that is saying is those models are equally plausible. So what I

would say to you is either we have a constant increase over time----

Q Yes.

A -- or in 2019----

Q (Inaudible) I think the previous page, July.

A -- the month, yeah, July 2019, there is sort of this step change increase.

Q So if we take all these various attempts using Dr Agrawal's data, where does it leave us? What does it tell us?

A Honestly, I'm going to go back to what I said before. It tells us we don't understand enough about ventilation. Do I think there should be, just my own opinion, there should be unventilated or nearly unventilated wards? Probably not, right? You don't want to close a building without ventilation. Does that seem to have an impact on Aspergillus? The data's suggesting maybe not. Maybe there's something else going on. Is Aspergillus the best choice? Maybe not, because COVID would look very different, presumably, and so I think that there is a lot of work to be done.

You know, and if you go then to the literature and other things to be done on actually what's proper ventilation, that does not answer your question about should hospitals not follow guidance. I think that's a loaded question that I will refrain from answering because I'm not sure anyone should not be following the guidelines that are there. I guess I would say how good are the guidelines and I'm not sure. The data suggests to me that it doesn't have an impact on something like Aspergillus but again, like I said, it would probably have an impact on something like COVID.

Q On that topic, I mean, I put this to you that Aspergillus was quite close to being too rare to do this work on.

A Potentially.

Q And rarer conditions like Cryptococcus, we wouldn't be able to do this on that, would we? You're shaking your head.

A Yeah, no. You need enough data and I think that's why Aspergillus and these patients, these patients particularly, for patients that get these types of infections get Aspergillus.

Q Yes.

A You to pick an infection that might be representative.

Q And it has to have enough cases?

A Yeah.

Q And this is quite close to the margin of not having enough cases?

A Mm-hmm.

Q Right. Now, I think we're almost at the end. The final questions before I see if my colleagues have any questions relate to the design of the whole exercise, and your role within it,

and your duty as an expert witness. So if we look at HAD Questionnaire 1, so bundle 44, volume 2, document 1, page 28. It's actually question 16. We asked you about the contemporaneous epidemiology, and I listed a series of pieces of work. These first three are by HPS on page 28 over the page. We have Ms Harvey-Wood, Mr Kennedy, Ms Harvey-Wood again, Dr Kennedy, Dr Kennedy and Jennifer Rodgers, and over the page, that might be the end of it.

In all cases, you responded, "It was not in our instructions to review and critique the management of infection control." These reports aren't about the management of infection control, are they? They're about epidemiologists like you trying to-- or people who are trying to be epidemiologists like you, trying to work out what's happened.

A Sure.

Q So why didn't you look at them?

A So that was not part of our instruction. Data was part of our instruction. Again, I will reiterate, I would have to go back-- that list that you have of the, what you call the "data dump," I'd have to go back and check that that's even in my folder. It seems like a long list of items that I'm not sure were in there.

Q Because even if it wasn't there, you know from your instruction letters there's a Public Inquiry.

A Yes.

Q Amongst other things, and so you must be able to assume or infer that there is a level of interest in this. So we go back and look at the letters, so that's-- I won't go to the letters, but would you have known about the existence of the independent review?

A So----

Q I realise you're in Washington on the other side of the world----

A Yeah.

Q -- but would you have known about the existence of the Independent Review?

A No.

Q The Oversight Board?

A No.

Q But you knew about the Case Note Review, albeit you didn't read it?

A Yes.

Q Because I'm wondering whether there's an obligation on an expert witness instructed for court proceedings, amongst other things, to enquire when something as significant as these instructions come through to produce a report that may be used in court proceedings to say, "You want me to design and deliver an epidemiological data study, has anyone else tried something? I might have a look at them first." Why wouldn't you want to do that?

A So it is my understanding, and I was brought in late and so I would ask that you confirm this with my colleagues----

Q Yes.

A -- that there was a desire for an independent, not having reviewed other items necessarily, assessment of the data. I was brought in just to look at the data and analyse the data. I was brought in rather late, so it was a lot of rapid work. Having been on the other side of the United States, I only became aware of the Inquiry, which I know my colleagues had been aware of previously when I started working on it.

Q Because I appreciate that not being encumbered by other people's biases and thoughts would be an advantage, but the counterfactual is that you just accept the data you're given, you fail to realise that there is a national database of occupied bed day data available from the NSS, you don't know why the water is said to be by some people contaminated or what the inadequacies of ventilation-- when that changed. Therefore, that being unencumbered by the biases of previous people actually results in a lack of value to your exercise, because it's not grounded in any form of reality. Now, I realise I've phrased that in an unfair manner, but is there anything in that you would accept?

A I think that being able to-- if, knowing what I know now, I did this exercise, I probably would have still started with an unbiased analysis and then, given all the time in the world which we didn't have, been able to read the relevant items and then ask questions about comparability and other things like that, and the reason that I point that out is because as I read items, and I did not read them all, there many many bundles, but I read items that largely you and your team pointed me to. I felt that there was a biased approach to begin with that water was to blame and we were going to somehow prove that there were no counterfactuals that I saw in the items that I was directed to. Now, that may not be a correct assessment of what was happening----

But you weren't directed to any items, the only document----

Q No, by you and your team.

But you were directed by your original instructing solicitors to the HPS summary review, which you didn't read, the Case Note Review report, which you didn't read.

A But I was not directed toward those.

Q Well, chapter four criticises the Case Note Review and you signed the report, so you had access to that, and it's

full of counterfactuals. Chapter eight is full of lots of things, and the HPS summary review, I don't think it's full of counterfactuals, but it's full of detail. So how do you-- one of the questions it's been suggested I ask is this, is that you've willfully ignored data in order to design the study. How do you respond to that?

A I would say that that's an unfair statement because I did not have access to those data. I was given a data set by an NHS entity on NHS data, so maybe I naively trusted them.

THE CHAIR: Sorry for asking you to go back. You used the word "biased" a moment or two ago. Could you just tease that out?

A Yeah, so I think when you approach a problem like this one, so for example, you're asked to model data and determine if you see any significant trends. If you are told that you're modelling data to show that water is a problem, that's a biased approach as opposed to saying, "All right, you want to know if there are trends there or not, I'll let the data tell me if there are trends there or not."

THE CHAIR: So I may have picked you up wrong. I thought at the stage you were instructed, you had understood that somebody else had adopted such a biased approach.

A We had understood that there was-- yes, an approach to say there was a belief that water, and this was our understanding, which I understand now is not what was being argued, that water is causing all of the infections.

THE CHAIR: And where did you get this information?

A This is from the discussions that I joined later in meetings with the CLO team that there was this idea that water was causing a huge amount of infections.

THE CHAIR: And when you say the CLO team, who do you mean?

A [redacted] is who we originally met with and, again, this could be my interpretation, but there was a lot of talk about a lot of-- this increase in infections being caused by the environment.

THE CHAIR: Was it specified beyond a lot of talk?

A Could you clarify?

THE CHAIR: I mean, who was doing the talking?

A So, I am not confident that I can recall exactly who was always at any meeting or saying there was----

THE CHAIR: No, who was----

A It wasn't me who was talking about that.

THE CHAIR: Bad question. Who, were you told, was adopting this biased

approach?

A Oh, you know, I'm not sure that that was clear. It was sort of an excited kind of idea around it, and I'm actually now not sure that that's clear who was adopting that approach but the idea was, you know-- So the first idea we had is, that's kind of crazy, because if you had dirty water-- say the water was terribly dirty, it's not going to cause all your infections, and so then-- And we never had the idea that it couldn't cause any, so we were really looking at, "Is this, you know, causing a significant problem?" and the extent of the dirtiness of the water was not clear either.

MR MACKINTOSH: A couple of things. I think I heard you say, but I may have misheard, that when we, as the Inquiry, supplied you with material, we didn't supply you with counterfactuals. Did you say that?

A I'm not sure I said that. I think that I said that I wanted to introduce the concept of counterfactuals.

Q We brought Jennifer Rodgers' statement and the CLABSI to your attention in the first questionnaire.

A Yes, absolutely. I don't believe I said you didn't give me counterfactuals because----

Q No, I may have misheard. The point I want to just tease out is that I think it's proper to put to you the position taken in-- I mean, I realise you had your letter of instruction, and you've explained to his Lordship what was put to you in meetings, and I'm wondering whether it might be appropriate to put this to you. What's your view of the suggestion, in fact, that none of the BSI that you have had in your dataset were caused by the water? Is that something you would agree with?

A So, I want to step back from the word "cause," just because I'm an epidemiologist.

Q Absolutely. Would you like to do "associated"?

A I'm going to replace it with "potentially attributable."

Q Okay.

A And I would say that I don't agree with that. I think that some of this is-- if, as there is evidence, there were growths in taps and things like that, it is probable that some of the cases, that is where they picked up that microorganism, or possible.

Q And I suppose the final question is to think about the intersection between statistical epidemiology which you've been describing, and everything else that Bradford Hill talked about, the rest of epidemiology. So, we've heard evidence from a lot of people who were involved in the hospital from before it was built, when the water was filled in 2013,

who managed the water system, doctors and nurses in the various departments who were in the Infection Control team, and given its dysfunction on both sides of the argument, in senior management, both on the clinical and management side, and we've reviewed the investigations carried out by the Health Board in 2018, by Health Protection Scotland that year and the following year. We've got access to all that material.

I do appreciate that you've been very careful to keep your answers tied to your data. I do know that Professor Hawkey will have to discuss whole genome sequencing, and Dr Agrawal will have to go ahead-- So that's what has to get passed, I think. But for the Inquiry, is there anything wrong with going back to Bradford Hill's postulates, looking at them all, and coming to a balanced conclusion about association and causation, which gives epidemiology its place but also gives experience, interventions, all these things their different place in the decision-making process? Is there anything wrong with that as a way of proceeding?

A So, there is nothing wrong with using it as he stated, as assistance in trying to decide. He was very clear that it doesn't prove cause, but rather it's guidance to epidemiologists. There are other models of guidance for causation that should be explored as well.

Q Right, but the point you're saying is it's not about proof, it's about association, it's about connection.

A Yeah. I mean----

Q You can't prove a connection with epidemiology. You can give a strong indication or a weak indication, but you can't prove it.

A I would agree 100 per cent with that. This is retrospective data. We know maybe the data quality isn't ideal. There are lots of questions about the data, and so it is very difficult to say that there is cause here. I don't think-- Without something like whole genome sequencing, showing clonal, you know, sort of transmission----

Q For which you might need appropriate samples and we can discuss that with people.

A Exactly. There is no way you're going to prove cause, but this is weighted evidence in epidemiology, right? Is there-- What-- How much do we think it is likely something's attributed? And I will go back to why I like to use different models, why I like different datasets around the same concept. Because if everything keeps pointing in the same direction, and this is a very Bradford Hill kind of concept, that suggests that it is more likely that that is true.

Q Well, thank you very much, Dr

Drumright. What I need to do, my Lord, is to see if any of my colleagues, and those watching remotely, have any further questions. I might, if we might, rise for 10 minutes.

THE CHAIR: If I may, I might ask one question, which is intended to be entirely open and not reflect on anything. It certainly is not intended to be open to any adverse inference. As Mr Mackintosh established at the beginning of your evidence, you signed a declaration which included stating that you understand that your duty in providing written reports and giving evidence is to help the Inquiry. Now, when you were instructed by CLO, how did you understand the word "help"?

A So, I understood that as giving an objective response based on the data, and what I said to the CLO was, "If I agree to participate, what I see is what I will show, regardless of whether or not that's popular," right? Because they are asking you to work as an independent witness. I know that when you have a setup like this, people might have their own opinions of what they want to see, but I'll only report on what I see in the data, regardless of what that is. So my idea of helping is objectively reporting what I can see in the data.

As things evolved and we, you know, came under Mr Mackintosh's team, I was more than happy, as Mr Mackintosh knows, to perform analyses, go after features. When I found out about other data I was very interested in, I requested, probably at his annoyance, datasets over and over. Because what I wanted to do was – now that I knew there was much more data – model that data in the same way I modeled mine to see if there were differences. So, very interested in sort of helping to understand what the data are telling us.

THE CHAIR: Thank you. Well, what we'll do now is take a break for, I would anticipate, no more than 10 minutes to allow Mr Mackintosh to check with the room whether there are any questions that need to be asked.

THE WITNESS: Great.

THE CHAIR: So can I invite you to return to your witness room?

THE WITNESS: Yeah.

(Short break)

THE CHAIR: Apparently one more question.

A Oh right.

MR MACKINTOSH: So what happened, Dr Drumright, is I forgot to ask this question, which relates to the NSS review, the Aspergillus data.

A Okay.

Q So if we can just go to the document, it's 44, volume 3, at page 222. So this is-- I think it doesn't necessarily

interact with you, but it might do. So when Ms Cairns responded to the HAD report, she didn't have access, because of timings, to Dr Agrawal's calculations. She then got them, and then she produced this document, and if we jump onto the next page, just to see-- It's not very long, and the executive summary is on page 1. I might just suggest that you quickly glance at that, because I want to ask whether you've reviewed this document and what your response is to it.

A So, I don't believe this is a document that I've seen yet.

Q Right. So, the point that is suggested, if you go to the next page, is that there seems to be a criticism advanced by NSS that Dr Agrawal's data doesn't fully meet the best standards of carrying out epidemiological analyses in terms of case definition and such things. If we go on to the next page, you see on paragraph 6 they discuss his choice of de-duplication process, where he appears to have de-duplicated by the admission episode rather than by 14 days or something.

A Well, yes, so----

Q What I was wanting to ask was, in a sense-- I appreciate that you didn't do his exercise, he did it, but if we assume for a moment that his approach is to some extent criticisable on the basis it doesn't de-duplicate in the way NSS would expect, or has a case definition that's perhaps a little bit softer than you might want, what effect does that have on the value of the numbers that he's producing?

A So, the first thing I will say-- Again, I will have you put this to him, but I work in the area of fungal infection, and of course with Aspergillus you cannot just rely on the lab data, and that's all he had, so the case definition they can criticise away, but if you don't have the complete set of information, you can't set up a case, and this was my argument about data should be standardised. Somebody decided on those cases or not, so what will happen is you will call more cases than are actually true Aspergillus cases in that instance, right? So he won't miss any, but he will get too many. And then in terms of de-duplication, I would ask them to go back and look and make sure that they think there is a 14-day criteria like there is for bloodstream infection.

Q I mean, I did wonder whether Aspergillus is different in that respect.

A It is very different, because treating a fungal infection takes a whole lot longer than that, and it's just far more complicated. I would be incredibly surprised if that was their definition and perplexed at who proposed it.

Q I mean, I can ask you this general question because it's something

I've been adopting as a principal and perhaps Dr Agrawal's work is the extreme example. From an epidemiological point of view, do these sort of criticisms of studies – case definitions being softer than they ought to be, de-duplication not being complete or not being entirely clear in people's minds – knock out studies completely or do you still have to use them a bit?

A No. So these are your limitations in a study and they limit the conclusions that you can draw, but they do not knock out a study.

Q I'm going to leave that there. I'm glancing over at my colleagues. Thank you, my Lord. I think, Dr Drumright, that's all the questions I have for you. Thank you for your assistance over the last few months.

THE CHAIR: Can I add my thanks, Dr Drumright? It is very clear that you have done an enormous amount of work, no doubt in preparing the report with your two colleagues, but also in engaging with the Inquiry, and while your evidence is only part of a great deal of other evidence that we've heard, it does appear to me that that process of engagement has been productive, and in addition to just thanking you for your whole contribution, can I highlight that? But you're now free to go, and I don't know if you're traveling far, but safe traveling.

A Thank you very much.

THE CHAIR: Now, I'm sure legal representatives require no reminder, but notwithstanding that, can I remind you that we'll be sitting at 9 o'clock?

MR MACKINTOSH: We won't, my Lord. We'll be sitting at 9.30.

THE CHAIR: Oh, it's 9.30?

MR MACKINTOSH: There was an outbreak of enthusiasm.

THE CHAIR: Well, I'm glad I raised the matter.

MR MACKINTOSH: There was an outbreak of enthusiasm in one of our notes, which did say 9 o'clock, but the important people in tomorrow's evidence, that is the witness and Mr Connal, think it's 9.30 and, in fact, 9.30 would be more than enough time to get through Ms Harvey-Wood's evidence before Dr Agrawal, and so it doesn't have to be 9 a.m. I'm sure that will be well received by my colleagues.

THE CHAIR: Right. And the timetabling is Ms Harvey-Wood at 9.30----

MR MACKINTOSH: Around an hour and a half, yes.

THE CHAIR: -- with targeting Dr Agrawal at 11 o'clock.

MR MACKINTOSH: Yes, that's the objective.

THE CHAIR: Well, I'm glad to have been corrected, and can I wish you a pleasant evening.

(Session ends)