

**THE COUNCIL FOR MEDICAL SCHEMES (CMS)**  
**INQUIRY UNDER SECTION 59 OF THE MEDICAL SCHEMES ACT**  
**HELD AT**  
**BLOCK A, ECO GLADES 2, OFFICE PARK, CENTURION**

**19 November 2019**

**DAY 14**

**PROCEEDINGS HELD ON 19 NOVEMBER 2019**

**ADV TEMBEKA NGCUKAITOBI:** Good morning everyone. We are continuing the Section 59 Investigation. We're starting slightly late today largely because of me. I got stuck in Namibia last night because SAA cancelled the flights so I could only fly this morning.

We have a statement to read before Dr Kimmie starts his evidence. It's largely a consequence of queries received from one of the schemes. The statement reads as follows;

10 Copies would be made available to everybody who is interested. The Panel would be hearing the evidence from the expert that is appointed by the Panel, Dr Zaid Kimmie.

Dr Kimmie has submitted a detailed report to the Panel. A copy of his report will not be made available today until the medical schemes and the administrators have had an opportunity to comment.

My understanding is that he will be using a PowerPoint presentation for his evidence today.

20 We received correspondence from Discovery Health on 12 November 2019 regarding Dr Kimmie's testimony. The request from Discovery was that they wanted an opportunity to comment on his report and check his facts before he reaches any conclusions or preliminary findings.

We considered the request and decided against it. It

is in the interest of fairness and transparency that Dr Kimmie should be allowed to present his evidence on the basis of the process that he and the Panel and the relevant administrators and the schemes have followed so far.

10 Most importantly the relevant administrators and schemes will have an opportunity to comment or challenge the findings and evidence reached by Dr Kimmie when they make their presentations and give their evidence sometime early next year.

The Panel will make available to Discovery Health to GEMS to Medscheme a complete copy of the report of Dr Kimmie containing his reasons and findings as well as other notes and data set relevant to each scheme and administrator. This will give the schemes and administrators a full and fair opportunity to comment. They will also have access to the transcript of today's hearing which will accompany the PowerPoint presentation.

20 Furthermore Discovery and GEMS and Medscheme have made various claims of confidentiality over information that they have provided to Dr Kimmie. We have asked Dr Kimmie to exclude this information from his presentations and his slides. We are accordingly respecting the request for confidentiality

in this session. But we will in due course engage Discover, GEMS and Medscheme regarding this information so that the transparency of this investigation is not compromised.

10 Finally, we do not want this investigation to be delayed. It is important that the parties involved continue to cooperate in a pragmatic way allowing it to be concluded as soon as reasonably but certainly by the beginning of next year. The intention is to have one more hearing in late January or February where the schemes and administrators are given an opportunity to present their evidence and to be heard. I may add that they are also entitled to submit written documentation to substantiate their respective arguments.

20 And thereafter the Panel will hope to write up the report and present it hopefully by March next a full report of the Panel will be made available. As I say the full copy of that statement would be made available. It is prepared by me and the members of the Panel so it is the report of the Panel.

Dr Kimmie, I should start by taking your oath. Are you happy with the oath or do you want an affirmation?

**DR ZAID KIMMIE**: Affirmation.

**ADV TEMBEKA NGCUKAITOBI**: Affirmation alright. Will you then say

after me, I and your full names.

**DR ZAID KIMMIE:** I Zaid Kimmie.

**ADV TEMBEKA NGCUKAITOBI:** Solemnly affirm

**DR ZAID KIMMIE:** Solemnly affirm

**ADV TEMBEKA NGCUKAITOBI:** That the evidence that I shall give

**DR ZAID KIMMIE:** That the evidence that I shall give

**ADV TEMBEKA NGCUKAITOBI:** Shall be the truth

**DR ZAID KIMMIE:** Shall be the truth

**ADV TEMBEKA NGCUKAITOBI:** The whole truth

10 **DR ZAID KIMMIE:** The whole truth

**ADV TEMBEKA NGCUKAITOBI:** And nothing but the truth

**DR ZAID KIMMIE:** And nothing but the truth

**ADV TEMBEKA NGCUKAITOBI:** Thank you. Can we just start with getting to understand who you are? Because you are coming as an expert and also an investigator but I think to substantiate your expertise we have a brief CV here but just tell us. Oh yes, page 34, yes we have your CV here at page 34 so maybe you can start by outlining your background and what qualifies you as an expert.

20 **DR ZAID KIMMIE:** Okay. Thank you. I have a PhD in Mathematics from the University of Cape Town and my undergraduate studies included courses in statics and economics. I also completed a Masters in Public Health at Harvard with a focus on bio-statistics, epidemiology and the design and analysis of experimental data.

Professionally I've worked as a statistician at the council for scientific and industrial research for about 6 years. I've also worked in

the social research sector as a researcher focusing again mainly on the analysis of quantitative data. And yes I think I hope sums up my professional background.

**ADV TEMBEKA NGCUKAITOBI:** What is your area of expertise?

**DR ZAID KIMMIE:** Statistics, statistical modelling, mathematical modelling and the analysis of data particularly survey data.

**ADV TEMBEKA NGCUKAITOBI:** And have you produced reports before that are similar, done investigations that are similar?

**DR ZAID KIMMIE:** I've done work for financial institutions that are very  
10 similar and this sort of data analysis and I've produced professionally more than 50 reports on a range of areas ranging from the analysis of election data to magnetic resonance imaging of underground water sources. So ja have- this is basically been what I do for a living is analyse data.

**ADV TEMBEKA NGCUKAITOBI:** And how long have you done it for a living?

**DR ZAID KIMMIE:** It gives away my age, somewhere close to 25 years.

**ADV TEMBEKA NGCUKAITOBI:** So your overall experience is 25 years?

20 **DR ZAID KIMMIE:** Yes.

**ADV TEMBEKA NGCUKAITOBI:** Plus your qualifications and plus your professional experience.

**DR ZAID KIMMIE:** Yes.

**ADV TEMBEKA NGCUKAITOBI:** Alright. Well, can you just tell us I mean I saw that in the report that you gave to us you've included some

academic literature and maybe you can start there in just telling us, what is the correct approach to be followed in this sort of analysis?

**DR ZAID KIMMIE**: Okay. Perhaps I could take a step back and identify the two questions that the Panel eventually settled on that needed to be answered and that I would play a particular role in assisting the Panel in coming to an answer. And I'm going to then jump through the presentation because I may go back then and deal with some of the process issues in more detail. But let me jump ... (intervenes)

**ADV TEMBEKA NGCUKAITOBI**: Ja feel free to present it as you see it  
10 fit.

**DR ZAID KIMMIE**: Ja okay. So in as since I was asked to try and answer two questions, these weren't necessarily the questions that were originally posed. The regional question was simply assist the Panel in understanding how the algorithms work and what data is being used by medical schemes to identify potential cases of fraud waste and abuse.

Based on some of the initial work I did on the methodological issues and some of the data we collected from the three parties, Discovery Health, GEMS and Medscheme. The Panel settled on two  
20 questions which was firstly, is there an explicit racial bias in the application of the algorithms that- and processes that identify cases of fraud waste and abuse?

And this I understood as in the application in design of these methods, is there a clear and evident link between the race of the provider and the manner in which the algorithm is implemented. And

that could happen either because it's an explicit statement that says something. Perhaps it's a parody, if black then do this or it could be based on something that was obviously and clearly associated with race. So that was the one question. Could we understand how that happened and if it happened?

The second question dealt not with the process itself but simply with the outcomes. In the outcomes of these fraud waste and abuse processes, was there a racial bias evident in the data? Okay so those are two questions and they require different approaches.

10           The first question I answered by reviewing the submissions that the various parties that made about their methods that they employed and by observing demonstrations of how these systems operated and what sort of outputs they produced. So in each case I reviewed data and I attended a session where the various teams from the three parties involved demonstrated how their system worked.

**ADV ADILA HASSIM:** These sessions you attended are they at their premises?

**DR ZAID KIMMIE:** Yes. I was- I attended each of their ... (intervenes)

**ADV TEMBEKA NGCUKAITOBI:** Look, can the gentleman there help us  
20 with the speakers because we're getting an interfering sound?

**DR ZAID KIMMIE:** Alright okay so.

**ADV ADILA HASSIM:** The question was- before you get into your answer to the first question, perhaps if you can just tell us what steps you took to engage the three parties and what type of engagement you had in order to acquire the information you needed in order to answer



the question.

**DR ZAID KIMMIE:** Right. As I said firstly, each of the parties had made a submission, an initial submission to the Panel in which they attempted to answer some of the questions that had been raised. I with the assistance of the Panel arranged a session at each of their premises at which their system was demonstrated to me in a live fashion. So someone logged on to the system, ran various queries and demonstrated what sort of output was used and the discussion would then focus on the details potentially of what algorithms were used and  
10 how the decisions were made about whether a case was investigated or not based on the output of these systems.

I should say that the application of the systems varies quite substantively between the three providers and in two of the cases the algorithms themselves are not available for inspection. So I could see the data but I was able to see what data the system used as an input. And based on that observation and on those interactions, I am convinced that there is no explicit racial bias in the application and implementation of the algorithms nor is there any explicit racial bias in the design of the algorithms and systems as they are currently  
20 implemented.

I should- that's fairly a straight forward statement but perhaps just to explain that in essence this means that the data that goes into these systems is something called a PCNS number, the practice number, and that practice number has no relation at all to the racial group or designation of the individual practitioner to whom that number

belongs. And so that really is the key component and secondly that the analysis within the system is only done on PCNS numbers and on discipline specific criteria. So no geographic data or other immediate proxy for race is used in that system.

So I'm fairly comfortable in saying that there is no explicit racial profiling.

**ADV KERRY WILLIAMS:** Just to clarify- can you clarify that statement just because it was rather broad when you first made it? So you said something like there isn't explicit racial bias in the application or  
10 implementation.

**DR ZAID KIMMIE:** Yes.

**ADV KERRY WILLIAMS:** What you understand then about implementation?

**DR ZAID KIMMIE:** Okay, so the design of the system doesn't include any racial data. And let me then qualify, when I say explicit I'm saying that there is no indicator or variable that identifies whether a provider is black or not black. And the way in which that system is then implemented and the decisions on made on which cases to select, no explicit use is made of the race even in cases when it may apparent  
20 what the race is. There appear to be decisions made on operational grounds and focus grounds to identify particular types of practices but in that decision there is no explicit decision point that says if black do this and if not black do that.

This does not however mean that any of the factors used and the manner in which they are implemented are to use a phrase "race

blind”, that’s and entirely different question. That question is one of outcomes rather than process.

**ADV TEMBEKA NGCUKAITOBI:** I still want to just deal with the first conclusion. I mean if you could just take us through the process that you followed before you came to those conclusions and the evidence that you analysed.

**DR ZAID KIMMIE:** Okay right. So as I said the first bit of evidence was the description of the systems provided by each of the parties and they were relatively detailed about what the system did and how it did  
10 it. Even in cases where the precise details were proprietary so where it was black box in a sense that people would put claims data in and receive an output, a score, a risk score. There was some description of what factors were used and there was a statement that only the data in the claims line, the number of the provider, identification number of the  
20 patient, the code that was used, the value, the date. That was the only data used in the operation of the system.

During the interviews with each of the parties, I then observed how the system operated and what outputs were produced. And it seemed clear to me that the description that was given was accurate  
20 that there wasn’t a point at which someone turned off the system and did something funny and then turned it back on again right. So the system only used claims line data in the generation of its output. And I was able to see that demonstrated by each of the parties.

**ADV KERRY WILLIAMS:** Just to ask a further question, and then how is- I mean without breaching any confidentiality, how is the output dealt

with?

**DR ZAID KIMMIE:** The output in all cases is a score and highest scores mean higher probability that some sort of deviation from normal behaviour occurs. So the standard operation of these systems is to choose a bunch of factors usually specific to a discipline to identify what the normal behaviour is within that discipline and then to score the behaviour of each provider relative to that normal score right.

**ADV KERRY WILLIAMS:** So in relation to all three, do they generate a score?

10 **DR ZAID KIMMIE:** They all generate a score. Sometimes it goes from 0 to 1000 and other times it may range from 0 to I guess 150 or thereabout.

**ADV KERRY WILLIAMS:** Okay.

**DR ZAID KIMMIE:** So but it's a score which says higher numbers are linked to potentially more deviation from the norm. And the idea behind these systems is that if you find the large deviations that will indicate irregular behaviour- potentially irregular behaviour, in all cases that score is not the final determinant, it is used as an input into an investigation process. So no decisions are ever made on the basis that  
20 you happen to score 999 out of a 1000.

**ADV ADILA HASSIM:** So that, is that then- you would have seen in the hearings thus far that the forensic investigators in the letters that they write to the health service providers, you'd have seen a reference to you are outside of your peer group.

**DR ZAID KIMMIE:** Yes.

**ADV ADILA HASSIM:** You are, ja. You exhibit your conduct is outside of the peer group. So that would be the peer group, the reference to the peer group.

**DR ZAID KIMMIE:** Yes.

**ADV ADILA HASSIM:** Is what you are referring to.

**DR ZAID KIMMIE:** Yes.

**ADV ADILA HASSIM:** As where the risk- that's the risk score.

**DR ZAID KIMMIE:** Ja. Okay so, to take a simple example, if your general practitioners use a particular code 20% of the time and one  
10 group general practitioners uses that code 90% of the time. If the system is designed to identify that as a potential source of fraud waste and abuse, the providers who use that code at those higher rates have risks scores.

But their scores are never calculated on a single factor, there are a bunch of factors that are all put together and based on the cumulated score you have high or low risk ratings as it were. And each case there is an investigation which proceeds after a case has potentially identified. I will say I don't think ... (intervenes)

**ADV TEMBEKA NGCUKAITOBI:** Explain this because I'm not sure if I  
20 understand it. So you have Dr X (inaudible) here and you put in the numbers and then it gives you a score that shows there's deviances.

**DR ZAID KIMMIE:** Yes.

**ADV TEMBEKA NGCUKAITOBI:** But you don't investigate simply because there's deviances. What are the steps that are following between the identification- okay which is a race neutral score.

**DR ZAID KIMMIE:** Mm.

**ADV TEMBEKA NGCUKAITOBI:** But then there's human involvement that ultimately leads to the decision to investigate. So between the score and the decision to investigate, what will the schemes actually do?

**DR ZAID KIMMIE:** Okay. I think the single most important thing to understand about all of these systems is that they are resource limited so none of the systems operates on the basis that a particular score will qualify you for an investigation. So there's no magic number that  
10 ever pops up that says if you happen to score more than 903, we have to investigate you.

There's an intent to investigate as much as possible. There are two sources of information generally, well three, there are tipoff's and hotlines and all schemes prioritise tipoff's and hotlines, those are given the first priority to resolve as queries. The remaining space, the remaining capacity is then taken up so the decision is always how much capacity do we have and so then how many cases can we take from these systems.

And the cases are then identified on the basis of which  
20 particular areas or disciplines or practices have potentially higher risks involved. So that decision about which cases go through is a combination of a risk score and a business decision about what is more important to be investigated. But those are made on the basis of industry knowledge about what sort of behaviour needs to be addressed.

So to put it plainly if you had a set of behaviours that were costing the scheme R12.50 a time and a set of behaviours that was costing the scheme R85 000 a time, the commercial decision is we'll investigate the cases with higher consequence first. Within that decision space the output from these risk rating tools are used to prioritise cases. So there's a business decision that says we will focus on discipline X and behaviour Y and the system will then generate the practices and the providers that have high scores within those parameters.

10           That's not- the behaviour varies quite a bit across schemes. Sometimes the business decision is the prime driver. The analytics tool is there as a backup to say once we've decided what we going to look at the analytics tool will kick in and give us cases. In other cases the analytics tool also generates cases that go into an investigation space. And I'll go back to that first point, the limiting factor is how much capacity is available to conduct investigations and that leads to rates of identification of fraud and waste behaviour.

20           So I think I'd be on safe ground to say that in any of these cases, any of the three providers if you double their investigative capacity, they would potentially double the number of cases that they identified.

**ADV TEMBEKA NGCUKAITOBI:** Sorry. So I mean I think it's the first time that we've heard that sometimes the primary driver can be a business rationality. So we've often got evidence that says oh well, we've got this blind algorithm and then it identifies (inaudible\_audio)

subject to the investigation and then we conduct.

Now when you say sometimes the primary factor is a business decision, how does that actually manifest? Do the scheme say well, we think that psychiatry is a risk area and we must therefore devote the energies and the resources into psychiatry?

**DR ZAID KIMMIE:** Okay. So it will be a combination of an actuary or type decision that someone perhaps identifies a type of behaviour. Classic example that was used (inaudible\_audio) some respiratory decease. A normal hospitalisation pattern is one day for a certain  
10 proportion of people suddenly large numbers of people are being hospitalised for three or five or seven days.

It's a potential risk because of the cost implications and that would then be identified as which providers are responsible and can we identify what sort of- where they are and should they be investigated. So that's- that would be a decision driven by an actuary or alert about what the potential cost to the system will be.

In other cases it may be generated by a clinical concern about a certain type of behaviour, perhaps the number of people receiving treatment for heart related decease in a particular area spikes or sales  
20 of particular medical devices spike. So with no clinical reason and then that would lead to the decision to say let's focus on that behaviour because it's more important.

So I would say that if the statement that it is a blind system that pops up numbers and that people just take them off the top whatever the highest numbers are, is not accurate. It is a combination



of a commercial decision about what is more important as well potentially decisions about what capacity is available to investigate particular types of cases.

**ADV ADILA HASSIM:** Is that true across the three schemes and administrators that you investigated?

**DR ZAID KIMMIE:** To varying degrees. In some instances there is more reliance on the tool to generate investigative options but within that restriction there is still a decision. It's never a decision that says it's Monday morning, the system is run, let's pick out the 300 top cases  
10 and hand them out to our investigators right. That's the decision is always going to be what are we going to focus on and what's important in terms of commercial cost and that is not an automated decision ever.

**ADV TEMBEKA NGCUKAITOBI:** And who makes these decisions?

**DR ZAID KIMMIE:** That also varies. There may be a team that reviews the data on instruction. They may have been asked to focus on a particular area. They will review the cases and decide which ones go forward. In other cases it may be up- they be more latitude granted to individual investigators to decide which cases they will take on if they have the space available.

20 And those decision are potentially made on the basis of their specialisation so if for example someone has a long track record and has been focussing on pharmacies and their capacity is now available to take on four cases, they will choose the four highest cases from pharmacies. They're not going to suddenly pick oncologists for example right.

So there's- so that is not consistent across. In some cases there may be more of a top down approach in the allocation to investigators. But even in those cases where there's a top down approach it will- that top down approach will be based on as I said capacity and the particular capacity available. It's not taken out of a hat and given to the next person in the queue. There is some business sense behind it.

**ADV ADILA HASSIM:** Ask you a question about, its double barrel question about tipoff's and hotlines. Of the number of the FW  
10 investigations that take place, what percentage of that is made up by tipoff and hotline reports? And the second question is related to that what you've just described as the decision making that goes into which cases to investigate does that also apply to the tipoff's and the hotline complaints?

**DR ZAID KIMMIE:** Okay so it varies by the scheme but I would say somewhere around 50% would be a good guess for the number that comes through via tipoff's and hotlines. There- on some cases there's a business decision that all such cases must be investigated within a set period of time. So that would be in- that isn't a- that's a sort of  
20 obligation, an internal process obligation. But those cases tend- will receive priority so they will all be investigated.

Tipoff's and the operation of the analytical systems aren't you know hermetically sealed off from each other though. Clearly if a tipoff is received about a particular practice and then a second and then a third tipoff, it would be derelict of the administrator not to look at the

totality of that provider and see if there was a pattern. But all of the tipoff's and hotlines are investigated because they're given priority.

**ADV TEMBEKA NGCUKAITOBI:** When you say look sometimes there's a business decision taken and the forensic investigators are instructed to focus on an area, where is that business decision taken?

**DR ZAID KIMMIE:** I'm not sure how the decision is taken across all schemes. I focused my discussion in detail with only one of the schemes. But that would be a case where a risk committee would meet, would be guided by the actuarial input that would say here is a  
10 potential risk. According to our calculations this set of claims should be 3% and now 3.5%, if it's left unchecked it may have very serious consequences.

So potentially this could involve millions of rands so let's look at what's happening there in particular. Is there something that analytics can tell us about where those cases are and that may then feed into an investigative process to identify where those cases would go?

**ADV KERRY WILLIAMS:** Is that the team scenario you were talking about where a team makes a decision?

20 **DR ZAID KIMMIE:** Yes.

**ADV KERRY WILLIAMS:** Okay. Also just to be clear, you speak schemes and administrators. Could you just clarify the three schemes and administrators you are talking about and just going forward how you are going to speak about them?

**DR ZAID KIMMIE:** Okay. So to save myself some time, Discovery

Health is an administrator and it has various schemes that it operates, Discovery in particular. MedScheme is a scheme administrator, it has various medical aids, PolMed I think, Bonitas etcetera. GEMS is the exception because it is a medical aid. So when I refer to GEMS as an administrator, that's not precise. GEMS have subcontracted-contracted Metropolitan Health to run their analytics so Metropolitan Health actually runs the data, does the identification and acts on instruction from a GEMS forum which will say look at this or identify that or give us more information or tell us who the top 30 providers are  
10 who have exhibited this particular type of behaviour.

Where are these hospitals that are admitting people with bronchitis for four days at significantly higher rates than we would have expected clinically. Okay so something may be happening, tell us who those providers are and let's then find a solution to the problem. And the solution may involve an investigation. It may involve a rewriting of the rules as of a heavy handed approach. The particular scheme you're on is no longer going to hospitalise you for more than one day for bronchitis. But that's I would say an extreme approach. Often it's finding out exactly where the problem is and stopping it would be the  
20 approach.

**ADV ADILA HASSIM:** If my colleague's last question is on this topic I'm interested to hear about your methodology in relation to- do you have anything more on the answer to question one first before-

**DR ZAID KIMMIE:** Ja I think the word explicit is there for a reason. As I said it's not- it does not necessarily mean that the system is race

blind with respect to outcomes. It's simply that race does not appear to factor in the decision, the key decision making practices but in any explicit form so it's not designed with an explicit racial basis in mind.

**ADV TEMBEKA NGCUKAITOBI:** The machine doesn't say black or white.

**DR ZAID KIMMIE:** Yes. Ja, let me give you perhaps a simple example. If you were in a South African context looking at candidates for a particular position and you decided that race would play no part so some third-party would take the CVs and cleanse them of any racial  
10 identifying information, the name of your high school, your name of your partner, your name etcetera, the university you studied at if that were an identifier. And then you would make the decision simply on the data in front of you.

You would not be surprised that you would appoint particular demographic groups because that would represent historical patterns which are evident in the outcomes even if they're not part of the process. So if after all of that you picked out the CVs and your three top candidates were white, you wouldn't exactly fall off your chair and say something has gone horribly wrong with our system for example.

20 **ADV TEMBEKA NGCUKAITOBI:** What I did want to understand about when you say the system is not, it doesn't explicitly- so you could have one of two possibilities. The one is that it's deliberately designed to suppress race, the other is that well, no one thought of race.

I mean have you looked into that or- we had evidence here from another expert that says in the South African context you can't be

race blind, you can't design systems that overlook race.

**DR ZAID KIMMIE:** Okay. All of the parties have indicated that they are unaware of the race of the providers because all they deal with is a provider number right. And I think you've had some discussions with some of your- the people who have given evidence about whether you could infer the race of someone from their name and that, I will get to that in more detail, it's something that you can clearly do in a South African context but if you don't do that then you have any information at all right.

10           So the information that was provided to me by each of these parties does not contain race information. So on that basis they would not be able to answer the question are the outcomes racially bias or not because there is no way of deciding on the data that they have what the result is.

**ADV ADILA HASSIM:** That is the outcome of algorithms ... (intervenes)

**DR ZAID KIMMIE:** Yes.

**ADV ADILA HASSIM:** The risk ... (intervenes)

**DR ZAID KIMMIE:** We talking about the outcomes as the totality of that process. Running the algorithms, identifying cases, identifying which  
20 ones are investigated, deciding on the base of investigation, whether there's an actual case fraud waste and abuse. The output of that machine is something that the providers- these parties have said they cannot answer the question based on that data because it contains no information about race.

**ADV TEMBEKA NGCUKAITOBI:** What I was trying to explore in it so

there are two things, well I can't give you information on race because my system doesn't have information on race. But there's another question, why does your system not have information about race?

**DR ZAID KIMMIE**: That I think is best answered by them.

**ADV TEMBEKA NGCUKAITOBI**: I appreciate that, ja. Alright look, if you are finished with topic one.

**DR ZAID KIMMIE**: Yes.

**ADV TEMBEKA NGCUKAITOBI**: And you have nothing else to add, you can move to topic two.

10 **DR ZAID KIMMIE**: Okay right.

**ADV ADILA HASSIM**: As you enter topic two, can you start off by explaining because now you are asked to answer a question that requires you to assign race.

**DR ZAID KIMMIE**: Yes.

**ADV ADILA HASSIM**: And you've already told us there isn't explicit race bias in relation to the first question. So before you begin, if you can explain how you conducted your research in the methodology in order to come to an answer on this question.

20 **DR ZAID KIMMIE**: Okay alright. So first off it would seem just on the basis of being a South African and having grown up in South Africa, it would seem on the face of it a reasonable assumption to make that you could assign race on the bases of someone's name. In fact, with something I did many years ago for a financial institution in a case where no race information was collected.

You can infer race in South Africa based on name. I would say

without any research or too much thinking about it and you could potentially infer someone's race from where they lived because we are still in many cases a sort of relatively racially homogenous society within geographic area. If you lived in Umlazi you'd be- you'd have a fighting chance of saying that person was black. It may not- there may be the odd case where you would fail but certainly you'd be better off just saying that they were black.

South Africa of course is because of our history has collected racially specific data in almost all cases. Census data, the Department  
10 of Labour demands that you racially specific data to in fulfilment of employment equity plans etcetera etcetera. So the question of whether you can infer someone's race is never something that occupies much thought in the South African context because we have access to racial data in many cases.

What I did do in this case was examine the academic literature to understand whether there were methods that were being applied, how common these methods were and how they were being applied if at all to assign race based on a name. And that's I guess the, what- okay I'll leave that.

20 Okay so the question is can we find a racial- classify using the data at our disposal. So let me just talk briefly about what that data is. The master data set that every provider has access to is the PCNS database which is maintained on behalf of the Council for Medical Schemes by the Board of Healthcare Funders. They have the system which includes the name and surname and identity number and practice



address and home address and financial information, bank account details and ID number of every medical service provider who's going to claim from a medical aid.

Okay, so in order to claim from a medical aid you need a PCNS number and the PCNS number is assigned by the Board of Healthcare Funders. That form and that data does not include any information on race. It includes all of that demographic information with some variations, I mean it's not a perfect database, and it contains a code that will indicate what your discipline is and what your sub-discipline is  
10 if appropriate. So your discipline is whether you're a GP and if you're specialist, whether you're an oncologist or cardiologist etcetera.

So that information is in the system and there's no information on race. There are two numbers involved, there's a practice number and a personal provider number. And the way the system at least currently operates is that every provider has their own personal number and sometimes it's the same as the number of your practice if you're a sole practitioner. And sometimes if you the first partner, it's the number- the practice number is your number.

But every individual has their own number. It is true that some  
20 individuals have two numbers or more because they run multiple practices and sometimes there are glitches in the system. But generally we can say if you're a provider you have a unique identifier and that's what the medical aids will use to identify you when they process the data on their system.

So they see just the provider number as a claim and they can

infer that additional demographic data afterwards, your name, surname etcetera. The name and surname you think would be an obvious thing to have but in many cases that information is not there. And that's particular the case for pharmacist who register a practice under a name so that potentially some historical behaviour and opticians.

The system also has a flag to show whether you're active or not so whether you're actually allowed to claim or whether your practice number has been deprecated. Okay so that's the PCNS database.

**ADV ADILA HASSIM:** So you're saying surnames?

10 **DR ZAID KIMMIE:** Surnames yes, trying to keep track of where I'm going. Okay so the question is can we use surnames to classify by race? If we look at international practice, the answer is yes. It's a fairly common procedure in cases where there is no racially specific data. Ideally of course you'll ask people to self identify. But in many data sets particularly in countries other than South Africa, people do not have an additional self identification.

Okay so the paper I'm going to use, it's one of the first ones that pops up Fiscella and Fremont and it's a review paper so it's useful in the sense that it list many other sources and it looks at the use of  
20 geo coding and surname analysis to estimate race ethnicity particular in a health environment. And as it turns out I'm going to have removed the follow up references from this paragraph because they list multiple papers which then list multiple papers as you would do if you're doing literature survey on this space.

The US Census Bureau has for more than 50 years been

running a Hispanic name databases to assign race to Hispanic people in the United States pretty successfully. It's a sort of as a begets the shining example of how these systems work and its proof to be more than 90% accurate. So 90% of the time they can classify someone as Hispanic based on their surname and if that's followed up with a personal identification and someone specifies what the actual self identified ethnicity is, they're right 90% of the time.

But there are- what it's useful for and it's been used to identified Arabic, South Asian and South East Asian populations and  
10 what they all have in common is that if you look at their names they are not Anglicised or Europeanised, Phon P-H-O-N, Po or Geab, those are not English names. And the idea is when these populations have distinctive naming practices and they are relatively homogenous then you have a decent chance of inferring race based on the use of that surname.

And so there are multiple examples in the literature where we using kind of Chinese, South Asian, there are a couple of examples in the UK that used Indian, Pakistani names. There are a couple of studies in the US that use Arabic names for particular populations, in  
20 Minnesota there's a large Arabic population for example.

So people have constructed databases of these names which can then be used to infer race simply by saying as a simple look up, if your name is Hassim, chances are that you're of South Asian decent and of you go. If your name is Williams, not Williams even if you and I do happen to know many Williams' who are not white. You are not

going to be able to use Williams to say whether someone is black or white. Ngcukaitobi should be fairly straight forward right.

The question is that's a heuristic that I can do but there's a lot of names so how are we going to do it for this particular purpose will be the question to ask. There's some data about how accurate these systems are, more than 90% accurate and the condition that seems to drive whether or not these systems work is homogeneity and distinct naming conventions.

Now I would argue that and we'll do a little- you can do a little  
10 sanity test on this yourself, I'll give you a chance. This certainly that is the case in South Africa if we're dealing with African, Moslem and Indian surnames. They are fairly distinctive. There are other groups in which you wouldn't have much of a fighting chance of deciding whether someone was black or not. According to our existing classification schemes right but for those three groups African, Moslem and Indian, I think it's a- it should be a reasonable straight forward task.

I mentioned this earlier. I've been told informally about cases where this has happened in South Africa for epidemiological studies but I've seen nothing in the literature where this has been used in South  
20 Africa. I also mentioned that I did such an exercise about 12 years ago for a financial institution. But that obviously was- the information didn't go beyond that particular research project.

Okay. So I'm going to get back to the ... (intervenes)

**ADV ADILA HASSIM:** Tell us what the size of the PCNS database was that you used. So you're now going to tell us about how you assigned

race on that.

**DR ZAID KIMMIE:** Yes.

**ADV ADILA HASSIM:** But can you tell us first, what was the number of?

**DR ZAID KIMMIE:** Ja. There are approximately 150 000 unique observations in that PCNS database, that's 150 000 practitioners. Not all of them are active and I think if I looked at the active list I think it was about 77 000 who were active between 2012 and 2019. So in any particular year the number of active practitioners can vary as people  
10 register on the system.

Just because you're on the system doesn't mean that you actually being paid by a medical aid to provide service to their member. Some people simply register on the system. I think there was a bit of a fuzz about government employees registering on the system about three or four years ago to supplement their salaries by doing private work and there were some issue about what those numbers were and what it meant.

But that's the sort of numbers of people we're talking about somewhere in the region of 150 000. The number of distinctive names  
20 is just over 30 000. There are lots of- ja, I don't think this is- Venter's. There are more Venter's than anyone else I believe in that database and a ton of Smith's etcetera so there are lots of repetitions of surnames so we're looking at about 30 000 surnames.

Okay was that ... (intervenes)

**ADV KERRY WILLIAMS:** Sorry, just to be clear. Are you trying to

classify 30 000 or 77 000?

**DR ZAID KIMMIE:** 30 000.

**ADV KERRY WILLIAMS:** 30 000. Why less, why not the 77?

**DR ZAID KIMMIE:** Duplicate surnames. So we're talking if you're going to classify by surname, you're looking at the number of unique surnames. I don't care that there are ... (intervenes)

**ADV ADILA HASSIM:** There might be 100 Smith's.

**DR ZAID KIMMIE:** 150 Williams', they all going to go in the same pot. We're not going to use the first name to make any further judgement  
10 call. It's just the surname. We're not going to look at whether the person lives in Khaylitsha or Sandton right. It's just the surname.

And part of the reason for that is if this were general sample of the population, geography would be a good tool to use but we're not talking about a general sample of the population. We're talking about medical professionals and it's likely that they are higher income earners and they live in different areas. You're not going to just make a simple assumption on area. So that's- we may use it as I talk about it later as a sort of sanity check but for the moment it's simply surname and 30 000 of them.

20 Okay so before I get into the methodology of it, let's go back to the question we're asking. We're asking about racially biased outcomes and what that means is we're talking about different rates at which events occur for black and not black populations. I'm going to use that phrase Not Black rather than white because technically we don't know that people are white but we do know that- I will argue we

do know they they're black or not black. Bare with me on that ...  
(intervenes)

**ADV TEMBEKA NGCUKAITOBI:** Black is Indian.

**DR ZAID KIMMIE:** Yes.

**ADV TEMBEKA NGCUKAITOBI:** And African.

**DR ZAID KIMMIE:** Ja

**ADV TEMBEKA NGCUKAITOBI:** And Arabic.

**DR ZAID KIMMIE:** Yes. Okay and I'll get back to the point about ...  
(intervenes)

10 **ADV TEMBEKA NGCUKAITOBI:** It says here Arabic.

**DR ZAID KIMMIE:** Yes.

**ADV TEMBEKA NGCUKAITOBI:** Did you mean Moslems?

**DR ZAID KIMMIE:** Moslems yes, Arabic surnames that have an Arabic origin for example. I'll pick some out of a list later on- minds gone blank on that score.

Okay that methodological issue we're asking for a difference between black and not black, the safest and most conservative route is to assume in the absence of any other information that a person is white okay. The default classification is white, if you don't have a  
20 surname, you're white alright. If ... (intervenes)

**ADV KERRY WILLIAMS:** When you say safest, what do you mean by that?

**DR ZAID KIMMIE:** If we are going to find an effect we'd want to make- be absolutely sure or as sure as possible that we are not artificially finding anything and if we classify – so let's take an extreme version, I

simply said that everyone is white. And then you asked me is there a difference between white and black? And I can say no, there is no difference because everyone is white, right? Right, so the more people we classify as white the less likely we are going to be – to find an effect. The assumption we have is that if there's a difference it's potentially going to be higher among black than white. So if we classify people as white we will dilute that effect. So that ...[intervenes]

**ADV KERRY WILLIAMS:** So on your part you're being very cautious  
10 and conservative in your methodology.

**DR ZAID KIMMIE:** Yes, ja. So we are not going to say anyone is black based on any other information. If there is a glitch in the system and there's a PCNS number on a provider database that is not in the PCNS that, you know, could be a timing issue, they're white. I don't have any information, they're white, right? No name, they're white.

So that means that I am making it more difficult to find any effect because – I won't go into the details where but we can – the non-differential classification. So if we're only deciding on the basis of names and we're defaulting to white that will mean that we dilute the  
20 effect, the difference between black and white and I'll get back maybe when we do one example of a table to show how that effect would work.

Okay, so there are two approaches. One is a top down and one is a bottom up approach and they have different levels of conservatism attached. So one way is to go and do a little search on the web and find databases or African names and Indian names and



Arabic names and these are fairly common. Clearly there are people out there with a lot of time on their hands and they put together websites that contain Zulu clan names and the surnames associated with those clan names and the list runs into hundreds and hundreds and hundreds of clan names. There have been academic studies published on Sotho clan names and the use of family names, etcetera, etcetera. So I've listed some of them there. So there are publicly available lists of African names – ja, publicly available list of Arabic names and there are publicly available lists of Indian names. The

10 Indian cases potentially interesting because the South African history online contains a thesis published I think in the 1950s by someone who investigated the ships that brought over indentured labourers and in an extraordinary labour of love someone typed out several thousands – I think it was up to 80 000 names, individual names and surnames, retyped them from ship manifests which carried Indian indentured labourers to South Africa. So that's quite useful because we just harvest those lists of surnames.

So that it is a top down approach, it says here's the PCNS data, I'm putting it to one side, I'm simply putting a list of names and

20 then using that set of names I'm going to assign race to the PCNS database and then we'll see with assigned race then in further steps down to the data we receive from the parties involved, the three parties involved.

Okay, 89 000 names in total, unique names. There are lots of duplicates of various lists. That list of names, that master list, is freely

available, so I can via the panel make that list available and anyone, if they wish, can then apply that set of names to the data at their disposal.

Right. The case we're looking at is slightly different from what happens when people construct these external databases and there are even more sophisticated techniques using data mining and names generating algorithms which generate more names for these databases, there are several published on how these algorithms might work. So rather than using brute force how you generate an Arabic surname that  
10 is distinct, for example, but we're not going to go there, we're simply going to take what we have available.

The second approach is to identify the race by inspection of the names themselves. That means if you go through, so it's a deleting process. You go through the list of 30 000 names and you delete all the names that aren't obviously black, they're not obviously African or Asian or Muslim. Coloured subpopulation in general is just – ja, that's white, you know, we're not going to classify September or Adams or anyone of that, they're all white. So the default – now you can say, yes, you know, you know a Botha that's black. I think he was a deputy  
20 minister in the first cabinet. For our purposes he's going to be white, right? Even if there's a black Botha, we don't care, the conservative approach is to say that Botha is black, that September is black, right? I mean, white. Thank you.

**ADV KERRY WILLIAMS:** So the point being you have significantly underestimated the number of black practitioners in the PCNS

database.

**DR ZAID KIMMIE**: Ja.

**ADV KERRY WILLIAMS**: Okay.

**DR ZAID KIMMIE**: So the process you go through is removed off that list, I had some people do that for me and then I went back over the list and removed some more. And then we end up with 11 000 names in total that we think are obviously black and we can then add that to that external list and that is the classification I used. So it's a two step process. Go from the outside, go from the inside, combine the two. I'll  
10 get back to this point. At now stage do I look at or even compute the race of people identified as fraud. Right? In this process there is a Chinese ball, the database with the PCNS numbers, people of practitioners who have been identified as FWA is kept in one spot and the classification is only done on the master list. So no one has a sense.

Is that process error free? No. People make mistakes. Everyone's going to make mistakes, we are going to slip in the odd name that isn't black but still slips onto the list. Okay, so there's some error involved in this process and I'll get back to that in a moment. So  
20 we're not – it's not a claim that every name on that list is indisputably black but I'll talk a bit about what – how we deal with that issue.

So we had a final database of 98 000 names, we then applied that to the PCNS data and identified people as black or not black. Okay. I think as it turned out some 32.5 percent of practitioners were black by that classification. If we used the more conservative

classification I think we ended up with 16% black.

My guess, my – well, firm conviction is that that underestimates the proportion of black practitioners, you know, a more sophisticated system or through further iteration we would improve the classification but we have what we have for this and was not black, right? So if there were two practice numbers, there's one case where someone married and they've got a new practice number – I mean, they've got a new name with the same practice number and the new name wasn't the same as the old name, I mean, the old name was  
10 Zuma and the new name was Botha, they were white. So all conflicts resolved on the basis of that default behaviour, right? If there's a problem you're white – I mean, not black. Okay.

**ADV KERRY WILLIAMS:** Sorry, would mind just repeating the final results in terms of percentages, what numbers were black and not black?

**DR ZAID KIMMIE:** Off that PCNS database I think it was 32.5% but the PCNS database is not the important factor here. I'll talk a bit more about what our universe actually is, right? The PCNS database is simply the place where names live and we expect that every provider  
20 with a PCNS number is on that list and in fact 99.9% of them are. There are some providers that are not on the list and that, I believe, is just a timing issue. They may have been added to the list after I was given the data.

**ADV KERRY WILLIAMS:** You go on to conduct your analysis on those providers who are paid by the schemes.

**DR ZAID KIMMIE**: Yes, I'll talk in a bit more detail about that. Right. Okay, here's just, you know, for your own edification. Each time I run the presentation it generates a new list of names for me. This is a list of names from the PCNS database and those are the ones classified black. You can run through that list, have a look and you're going to find one or maybe two cases where you'll dispute the classification of every list of a hundred. I've run that process about ten times, just as a sanity check.

I should indicate that this is a sanity check to see what the  
10 error rate of that classification is, it's not a find names and then  
change your classification. Classification is done once, it's done. You  
could potentially go back and refine it and refine it but I'm going to  
stick with what we've got. This is just to check what sort of errors pop  
up. So okay, I haven't – ja, Khan, Moodley, Neti, Moagi. Ja, I can't  
find anything wrong with that list. Your mileage may vary according to  
your personal tastes. I'll get back to that issue.

**ADV KERRY WILLIAMS**: Is it Nicolas?

**DR ZAID KIMMIE**: Yes, where is it?

**ADV KERRY WILLIAMS**: Second line, first name.

20 **DR ZAID KIMMIE**: Yes, okay, somehow that Indian manifest had a  
Nicolas. That's an error, right? So I have run this at least ten times  
and I'm finding one or at most two errors per list. Okay? So we think  
somewhere around 99% of the time we're classifying as black people  
who are actually black.

This doesn't mean there is possibly a change that Sibiya on

this list or Maurimede is in fact a young white woman in Ventersdorp but I'm going to live with that probability, ja, because that is really unlikely. Okay.

What are the potential pitfalls that we have to avoid? We have to avoid differential misclassification, right? Differential misclassification happens when you cheat and cheating, in this case, would be saying the extreme example of cheating would be to say let me look at the list of people who are identified as cases of fraud, classify that list, do some work on the others but treat that as the main.

10 You are then putting your thumb on the scale because you're classifying outcomes and not input.

If you make a mistake, it seems fairly obvious that your name within the parameters we're thinking about does not affect whether you have been identified as an FWA case or not, like whether your name is Smith or Jones doesn't – has nothing to do whether you've committed fraud or have been identified as an FWA case.

There is potentially a relationship we know between whether you name is Smith or Mokoena, and that's what we're thinking about, but within Smith and Jones, that's a non-differential classification error  
20 and the same with Mokoena and Mofokeng.

So when we say a name is black or white, when we're making that error, it's a non-differential classification error, we are not peeking at the result and making a decision about whether you classified black or white.

Okay, so then we have a classification. I believe that it's a

reasonable proxy for race and I'll talk a bit about how we can test whether the reasonableness of that proxy actually influences the results we find. Right? That's something you have to do and I have done and I'll report on what the effect of misclassification potentially is. That complete list of names is also available for inspection.

Okay, any questions on the name stuff? Right. Okay, the actual statistical tests I'm going to use come from two sources, both of which are considered, it's actually bibles in their field for categorical analysis, Agresti, the 2<sup>nd</sup> edition and then for some of the inferential statistics, Rothman and Greenland's on Modern Epidemiology. Those  
10 are two very standard textbooks, the additions I have are slightly out of date but nothing – the changes ever happened at the margins rather than in the stuff that I'm going to be dealing with.

Okay, the first thing I'm going to do is look at the combined data. There's data from Discovery Health, GEMS and Medscheme for the period January 2012 through to June 2019. I will also make available the file I used to process the data and there are a couple of things I want to say about that data that's not on this slide.

One is that the data analysis was conducted in a package  
20 called – programme called R, which is the *de facto* standard data analysis package for statisticians. Right, if you are a working statistician 15 years ago, a proper working statistician, you used a program called SAS, which is, you know, a big multinational and 15 years later, if you're a proper working statistician, you use R, right? It's an open source package and it's a sort of poster child for open

source software, I would guess.

The data that's inputted from – that's provided by each of the schemes is imported in its original form as an Excel data sheet and the code that's used to conduct that analysis is available. So I have never opened the file that contains the Excel data, I have left it untouched in the form in which it was provided to me. All the analysis, every step, every decision is documented. If I decided, for example, a particular provider has included in their list of provider codes numbers that don't match PCNS numbers, you know, it will say in the data file this has  
10 happened, it involves so many cases and I have removed them from the analysis, for example, right?

Each of the providers gave me two data sets, a list of all of the people they paid and - a list of all PCNS numbers that had been paid for services rendered to scheme members and a list of PCNS numbers that had been identified as FWA cases. Those are two separate lists and I have maintained them as separate data sets when I do the analysis.

Over that seven and half a year period we have about 65 000 providers on that database.

20 **ADV TEMBEKA NGCUKAITOBI:** Medscheme didn't give you for 2012 you said?

**DR ZAID KIMMIE:** Yes, Medscheme didn't have the data available for 2012 so I just used 2013 data as a proxy for 2012.

The list of cases that were identified was fraud, waste or abuse consisted of 16 453 numbers. So these are all unique numbers,



right? So over that seven and half year period there are 65 000 unique PCNS numbers and over that period there are 16 500 cases of FWA and they're unique cases. So if you have identified as an FWA case by one provider in one year and then another provider in another year you only count once, you've only been identified – you're a case, right?

If you've been identified in multiple years by the same provider you're just one case, right? Okay. So it's just – it's a simple question. Over the seven and half year period were you identified by any of the providers as an FWA case? That's the question you ask. And about  
10 30.5% of all providers are black by our classification scheme. That's the table over the period of race and FWA outcomes.

Now I'm going to go through this in a lot of detail because I want to deal with some of the methodological issues and then hopefully we can deal with the other tables in a speedy fashion. Okay, so those are the numbers. The black, not black is what we call the independent variable, it's what we've structured in the dataset. The FWA, not FWA is the dependent or the output variable, it's the outcome of this process and those are the various numbers involved, 6 300 black, 13 500 black FWA, 13 500 not FWA, etcetera. That's your standard two by two table.

20 We are going to use something called a risk rate. That's the proportion within black of people who were identified as FWA. It's a simple calculation, we divide the 6 300 by the 19 903 and you get 31.7%. That's the risk, if you were a black practitioner over the seven and a half year period that you'd be identified as FWA. And you can calculate – ja?

**ADV KERRY WILLIAMS**: You will explain to us the distinction and the difference between risk rate and risk ratio.

**DR ZAID KIMMIE**: Yes, that will come in a second. Okay, so the risk rate you can calculate for black, you can calculate for not black and you can calculate for the population. So if you were looking at that you'd say the risk rate for black is 31.7%, the risk rate for the not black group is 22.3% and for the population it's 25.2%. All good and well, that's just what the data says, right?

The statistic we're going to use to decide whether or not there's bias is something called a risk ratio. That's what happens when you divide one risk rate by the other risk rate. So in this case we're dividing 31.7 by 22.3, that gives us a number of 1.42. That's the risk ratio and that tells us how much more likely you are to be an FWA case if you're black and if you're not black and the 1.42 you can read as you're 1.42 times more likely or you're 42% more likely to be an FWA case if you're black rather than if you're white.

So that's just the numbers. And now the table I've put in low percentage so you can see the details with all the calculations. So we have a risk ratio. The risk ratio, does it tell us if there's bias? No, it doesn't. It tells us that potentially something is happening in the table but we can't decide based on a risk ratio if something has gone – as if something unusual has happened, it's just a number at this point, right?

Right, the second thing we're going to ...[intervenes]

**ADV KERRY WILLIAMS**: Can you just explain why it's just a number, please?

**ADV TEMBEKA NGCUKAITOBI:** Why you say that it doesn't show bias.

**DR ZAID KIMMIE:** Alright. The question is ...[intervenes]

**ADV KERRY WILLIAMS:** I just want to piggyback on that. So your answer is well, because if something is more likely to happen to one group than another, doesn't that tell us something?

**DR ZAID KIMMIE:** Yes, it does. Okay, if – let's take a simple example – maybe not – you have a hundred coins, you know, or you're going to toss the same coin a hundred times. You know that each time you toss it there's a 50% chance you get heads, 50% chance tails. Then you sit  
10 and you actually do the hundred tosses and it turns out that you had 55 heads and 45 tails. Is the coin biased? There's something wrong with that coin, you know? And what if it was 58/42? Is that expected? What if it was 40/60? Is that unusual? At what point do you start to say this is just way out of line, something's wrong, I should not be getting – let's assume that you tossed that coin a hundred times and you got 99 tails? At that point you'd say this is really – this is egregious, something has gone wrong, someone is playing the fool. The actual statistical probability that you would get 99 heads if you tossed the coin a hundred times, you can think of it as you could do  
20 that once a second, every second since the start of the universe up to today and you'd be pretty unlikely to get a hundred, 99 tails. You could do that forever and not get your 99 tails.

So the question of why you can't simply say – you have to be able to judge ...[intervenes]

**ADV KERRY WILLIAMS:** What tells us that? Or where do we – okay.

**DR ZAID KIMMIE:** Okay for this two by two table we're going to calculate something called a chi-squared P value. I'm not going to go into the details of it, the standard test for two by two table is to calculate the statistic and it gives you a number and the number is somewhere between zero and 1 and smaller the number is, the less likely or the more unusual it is that that table is there by chance. Right?

So in this case – I mean I'll leave aside the details, the P value is two – each of the minus 142. That scientific notation, that  
10 means because we can't write that number, it means that it's a zero followed by 142 zeros followed by 2. Okay, so it's a long, long, long list of zeros.

**ADV KERRY WILLIAMS:** Which number? Oh, 26.

**DR ZAID KIMMIE:** On the right hand side, 26. Okay, so that P value is very small and the way we interpret a P value is ...[intervenes]

**ADV KERRY WILLIAMS:** Sorry, can you just do the numeric equivalent of that P value or is impossible?

**DR ZAID KIMMIE:** The numeric equivalent is a zero followed by a long line at 140 zeros than a 2, right? So it's a very small number. A P  
20 value closer to 1 is an indication that there's nothing unusual about the table, a P value closer to zero means something's unusual about the table.

**ADV KERRY WILLIAMS:** Sorry, not about the table, about that number, 1.42, right?

**DR ZAID KIMMIE:** No, about the table itself.

**ADV KERRY WILLIAMS:** About the table, okay.

**DR ZAID KIMMIE:** The 1.42 measure is one measure of what's going on with the table, the chi-squared P value just measures the table itself in totality.

Okay, so now this may not be, you know, entirely necessary but I'm going to give you a short lecture on P values. P values are one of the most poorly understood things in the scientific literature.

**ADV KERRY WILLIAMS:** Sorry, Doctor, but you will come back to these tables to ...[intervenes]

10 **DR ZAID KIMMIE:** Yes, I'll come back to the table in a second. So there have been journals, the journal of psychology, I think that has banned the use of P values, said you're not – if you put a P value in your publication, we're not going to accept it, it will be automatically rejected because there's – you know, what does a P value mean? People I misunderstand it, it's badly taught. If you've done your first year of statistics course, someone has told you that P values less than 0.05 are significant, that's in fact – it's a lie, it's falsehood but that's the way it's taught and so people will automatically apply a cut off 0.05 and go around waving their hands and saying this is significant, this is  
20 not significant. It doesn't mean that.

The technical definition of a P value is the probability that the table that you see has occurred under the assumption, the Null hypothesis assumption, and the Null hypothesis is the assumption of innocence [indistinct] that in this table Black and White and not-Black populations have the same risk of being FWA cases. So our baseline

assumption is that there is no difference between the Black and not-Black populations. If we assume that that is true and we see the data, the p-value tells us what the probability is that that data has occurred, given that assumption. So, whether it is unlikely or not.

So here is some sense of scale so you can put those numbers into some practical use. If you have a number followed by 10 zeros, that is the number of stars in the Milky Way. One followed by 13 zeros is how big our Solar System is. One followed by 17 zeros is the age in seconds of the universe. One followed by 22 zeros is the number of  
10 stars in the observable universe and one followed by 80 zeros is the total number of particles in the universe. Right? So, just so you can get a sense, one followed by 142 zeros is proportionately about twice the number of particles in the universe. It is a very, very small number.

You can think of it this way, it is approximately the same chance that you went out and you bought yourself a puzzle of 100 pieces and you took it home and you gave it a good shake and then you flung it up in the air and all of the pieces landed perfectly into a made up puzzle. It is that sort of likelihood. So the likelihood that this table has occurred in a population where there is no relationship between  
20 Black and not-Black and FWA is for all practical purposes, zero. It is impossible.

**ADV ADILA HASSIM:** So it rejects the Null hypothesis?

**DR ZAID KIMMIE:** Yes, there is a relationship in this case. It is a meaningful relationship and the scale of that relationship is measured by the risk ratio. We can take that risk ratio as meaningful.

**ADV TEMBEKA NGCUKAITOBI:** Sorry, I just want ... [Intervention]

**ADV KERRY WILLIAMS:** Please explain what you mean by that?

**ADV TEMBEKA NGCUKAITOBI:** Yes, specifically, I want to understand, if I look at this thing and you are a Black doctor and there are probably Black doctors in the room today, and they say listen, the probability is 31.7 for Blacks likely to be, the rate of 31.7 for Blacks likely to be FWA and the not-Black, I mean all of these are the conservative numbers that you have produced.

**DR ZAID KIMMIE:** *Ja.*

10 **ADV TEMBEKA NGCUKAITOBI:** Not-Black is 22.3%. That gives you 9% differential between those two. Now, is this natural? Did it just happen by accident? What is your statistical mind telling you about these numbers?

**DR ZAID KIMMIE:** That it is a logical, that it is an impossibility that this would have happened simply by chance. That number  $2E$  to the minus 142 tells us that the probability that this has happened by chance is unimaginably small.

**ADV ADILA HASSIM:** The equivalent of a jigsaw puzzle falling down all made up?

20 **DR ZAID KIMMIE:** *Ja.*

**ADV TEMBEKA NGCUKAITOBI:** So that is the relevance of the p-value?

**DR ZAID KIMMIE:** Yes.

**ADV TEMBEKA NGCUKAITOBI:** It is to explain the differential of that 9% between the Black and the White.

**DR ZAID KIMMIE:** Yes.

**ADV TEMBEKA NGCUKAITOBI:** To say that it is impossible that this is by chance?

**DR ZAID KIMMIE:** Yes, this did not happen by chance. I will talk a bit more about how we interpret that in a moment but, ...

**ADV KERRY WILLIAMS:** So equally, just to clarify this. So Black practitioners would be 42% more likely ...

**DR ZAID KIMMIE:** Yes.

**ADV KERRY WILLIAMS:** To be ...

10 **DR ZAID KIMMIE:** Identified.

**ADV KERRY WILLIAMS:** Identified as FWA and the point is this could not happen by chance?

**DR ZAID KIMMIE:** *Ja.*

**ADV KERRY WILLIAMS:** It is absolutely improbable that this can happen by chance.

20 **DR ZAID KIMMIE:** Yes. Okay. Now I am going to just give the caveat that when a result is meaningful is not the same as it being statistically significant. Right? So we are going to generally judge things based on three numbers. We are going to judge it on the risk rate for the population. We are going to judge it on the risk ratio which we have calculated and the p-value. So for example, you have maybe seen these studies and they happen on the front page of a newspaper where someone says, scientists have conducted a study and drinking coffee is linked to brain cancer. Drinking coffee increases your risk of brain cancer and it is a significant result. So they have calculated a p-value



and it is very small and they say it is a significant result. What they would not tell you is that the risk of getting that particular cancer is maybe one in ten thousand, one in a million, and if you drink coffee it is increased to three in a million.

So it would be a case of, for example, if you go out walking this evening in Johannesburg your chance of being hit by lightning is, let us say for argument's sake, one in ten million. If I conducted a study then said, if you go out walking this evening in Johannesburg wearing a green T shirt you will triple your chances of being hit by lightning, and it is a statistically significant result, if you were a rational person you would look at it and say, eh, one in three million? I am putting on my Springbok jersey and going for a walk. Right? So the base rate is something important. If the base rate is very small then an increase in risk is not meaningful. Okay?

The risk ratio itself must be different, meaningfully different from one. So if I said your, or conducted a study and the risk of being hit by lightning ... [intervenes]

**ADV TEMBEKA NGCUKAITOBI:** Just to take you back to the, um, so the data that you used, how many FWA cases from Discovery, from Medscheme and from the ...

**DR ZAID KIMMIE:** No, I am not answering that question. [laughs]

**ADV TEMBEKA NGCUKAITOBI:** Alright, is that part of your confidential undertaking?

**DR ZAID KIMMIE:** *Ja.* I am reporting on aggregate data at this point.

**ADV TEMBEKA NGCUKAITOBI:** Okay.

**DR ZAID KIMMIE:** So the risk ratio must be meaningful. One point four, a 40% increase is meaningful but you could have a statistically significant risk ratio which had at an increase in risk by 0.1%. So the number means something. Right? And then the p-value tells you, together with those other two numbers whether you can apply your sense to the statistic and say that it is meaningful. So in our example we have a base risk of 25%. That is pretty high. One quarter of the population gets FWA. We have a risk ratio of 1.4 which is pretty high – 40% increase. And we have a p-value that is very low. Altogether that  
10 tells us, collectively, that something meaningful is happening with this table. And therefore we can conclude that there is strong evidence that there is a racial bias that exists with respect to FWA outcomes. And that Black providers are 40% more likely to be identified as FWA cases than their not-Black counterparts. So that is a formal statement based on those three numbers.

There is one alternative measure of effect that we could use. It is not a standard use because these things that we calculate the risk ratio, they are very easy to turn into mathematical models and estimation models. We can say what is the increase in the number of  
20 Black FWA cases over the time period that could be attributed to this racial bias? And we could argue in this case that over the seven and a half years, the scale of the bias – another way of thinking about that 40% - is that approximately 1300 additional Black FWA cases have occurred over what would be expected if there were no relationship between Black and not-Black.

What I did, just to demonstrate how this would work, is I created a new race variable called, ...[Intervenues]

**ADV KERRY WILLIAMS:** Sorry, before you go there, just because you are moving quite fast, where does this alternate measure of effect come from?

**DR ZAID KIMMIE:** It is a calculation. Essentially what you do in this case is to say the base population rate is 25%. Right? So that is the number of cases identified. If there were no difference between Black and not-Black populations you would expect, on average, somewhere  
10 around 25% of Black practices to be FWA and somewhere around 25% of White practices to be FWA. Right? Because, the population and your two groups should not have different results because there is nothing racial about it. Under that assumption, 25% of nineteen thousand, it is a reasonably easy calculation in this case, is 5000, right, it is one quarter of twenty thousand. It is about 5000. The difference between 5000 and 6314 is the difference between the observed and the expected value and that is the quantification of what the scale of the difference is. So you could do that.

**ADV ADILA HASSIM:** ... [indistinct] is like the risk ratio. It does not  
20 say anything beyond that that is what the number says. It does not tell us anything about the significance or, ...

**DR ZAID KIMMIE:** No, no, remember that we can use it in addition to or in place of the risk ratio. So that number is a measure of effect size. The p-value and the base risk will tell us whether it is a meaningful measurement and in that case, it is. So you can say with some degree

of confidence that over the seven and a half year period the scale of the deviation can be measured either as 140% more likely, if you were trying to explain to someone or you can say the absolute effect has been somewhere in the order of 1300 additional Black FWA cases.

Okay, this is just an additional ...[intervenes]

**ADV TEMBEKA NGCUKAITOBI:** Can you just tell me why the White, the non-Black is actually three percent lower than the base risk?

**DR ZAID KIMMIE:** Okay, because this is just a balancing of numbers. In order to get the base risk you take the average of the two, the  
10 weighted average, the one has to be low, one has to be high. In these tables, once you have got the totals sorted, you pick one number and the rest are forced. It is a one degree of freedom on a table but you do not have to, I would ignore that statement actually.

**ADV KERRY WILLIAMS:** So, I am not quite sure I understand the significance of this slide still. So what? Or are you getting there?

**DR ZAID KIMMIE:** This slide?

**ADV KERRY WILLIAMS:** Ja. Sorry this slide I am on, not that one, not that one, sorry, I am looking at a different one – The alternative measure of effect.

20 **DR ZAID KIMMIE:** Yes. Yes. Okay.

**ADV ADILA HASSIM:** It is the same thing. It comes to the same thing.

**DR ZAID KIMMIE:** Exactly the same.

**ADV KERRY WILLIAMS:** So it is a check that you have done on your, for yourself.

**DR ZAID KIMMIE:** So if you were to say, there is a racial bias in FWA outcomes and that racial bias is extremely unlikely to have occurred by chance, so we think that there is some mechanism behind it. I will talk a bit more about what that means because at this point we are saying there is a clear correlation between the two. Right?

**ADV TEMBEKA NGCUKAITOBI:** I mean I do want to just come back to these two conclusions. The one conclusion being that we have got 42% greater chance of a Black person being identified for FWA and the other being that the differential of 9%, it is impossible that that occurred as a  
10 consequence of some accident? Now, I accept that from a statistical analysis you have done the exercise, I suppose it is a point related to Miss William's argument which is, have you then tracked what is the explanation for the 9% differential and the 42% likelihood of being an FWA?

**DR ZAID KIMMIE:** *Ja.* I think there is, that is a different question, the question of why that has occurred? What, statistically I can report on is a correlation between two things; being Black and being FWA. How that mechanism has worked is something that ...[intervention]

**ADV TEMBEKA NGCUKAITOBI:** I mean you can report on the third one  
20 as well which is that it is improbable that this occurred as a result of neutral factors or chance.

**DR ZAID KIMMIE:** There is something happening. It may be that there is some other factor that is linked to race and linked to the outcome that would explain it. But I have no sight of that other factor. Right? The Null hypothesis in that case, when you were examining

alternative factors, the Null hypothesis would be that it is linked to race. You would have to find evidence to the contrary. So this says there is this correlation. If you were to explain it away you would have to find the additional intermediate factors and make an argument about why those factors lead to this outcome.

**ADV TEMBEKA NGCUKAITOBI:** Won't you be asking a different question which is, my understanding of your conclusion is that there is a clear correlation

**DR ZAID KIMMIE:** Yes.

- 10 **ADV TEMBEKA NGCUKAITOBI:** In relation to race, so anyone who says it is not race must then produce their evidence?

**DR ZAID KIMMIE:** Yes.

**ADV TEMBEKA NGCUKAITOBI:** It is not as if there is evidence to establish why because you have established the 'why'.

**DR ZAID KIMMIE:** Yes. You have established that it happens.

**ADV TEMBEKA NGCUKAITOBI:** And what you have established is the correlation.

**DR ZAID KIMMIE:** Yes. *Ja.* Yes, okay. That is what the ...  
[Intervention]

- 20 **ADV ADILA HASSIM:** Okay, I will take it one step at a time so that is fine, we are good with that for now.

**DR ZAID KIMMIE:** Okay. Here is another statistic, now let me say at this point that data analysis is like a treadmill, you can just carry on. You can carry on finding more and more interesting things to analyse and look at, so I am going to focus here on answering just this

question, whether there is a racial bias and how it manifests itself but this is something that is worth further investigation. Over this period, if you were identified by all three providers, if you appeared on everyone's data base of FWA cases, the probability, you are 400 times more likely to be Black than White, than not-Black.

**ADV ADILA HASSIM:** You have to explain that again.

**DR ZAID KIMMIE:** *Ja.* So within those numbers, the risk rate, if you are identified by all schemes, the risk rate is four times higher for Black than it is for White.

10 **ADV TEMBEKA NGCUKAITOBI:** I don't think that is what the statement says here because it says over the period 2012 to June 2019 Black providers were 400%, forget about what is in the brackets, more likely to be identified by all three parties whereas your point now is, if you are identified by all three parties you are 400% more likely to be investigated.

**DR ZAID KIMMIE:** Yes.

**ADV TEMBEKA NGCUKAITOBI:** But here this presupposes, this is about you are likely to be identified, which is the slide we looked at earlier?

20 **DR ZAID KIMMIE:** Yes. So the outcome that we are looking at is we are looking only at those people who were identified by all three and within that group, you are four times more likely, the rate was four times higher for Black than for ...[Intervenes]

**ADV TEMBEKA NGCUKAITOBI:** No, no, I accept that. That is what your oral testimony is but I am trying to connect it with what is written here. It seems here ...[intervenes]

**DR ZAID KIMMIE:** Maybe it is a lit bit ...

**ADV TEMBEKA NGCUKAITOBI:** What you are saying is that you are likely to be identified as opposed to you are likely to be investigated if identified.

**DR ZAID KIMMIE:** No, okay, I will correct that. That is not what I meant.

10 **ADV ADILA HASSIM:** Sorry, so what did you mean?

**DR ZAID KIMMIE:** I mean within ...

**ADV ADILA HASSIM:** Is it as you have written it?

**DR ZAID KIMMIE:** No, as I think the writing suffered here. Within the group that have been identified by all three, there are 800 such cases I believe, 618 cases. The risk that you are identified by all three is four times higher if you are Black than not-Black. It turns out in this case that, you know, I think more than close to 70% of all of those cases are Black that were identified by all three providers. It is a very severe deviation from what you would expect.

20 **ADV TEMBEKA NGCUKAITOBI:** You have a chance of survival if you are White despite being identified by all three. That is your point, if you are non-Black as you call it.

**DR ZAID KIMMIE:** Yes. Okay.

**ADV KERRY WILLIAMS:** Just practically can you please correct that slide so that we are sure to circulate the correct version?



**DR ZAID KIMMIE:** Okay. So this is what I talked about earlier on.

**ADV TEMBEKA NGCUKAITOBI:** I think we should probably take a break. It is five minutes to one and then I am sure you will still continue maybe for another hour when you return. I do not know how long you still have to go.

**DR ZAID KIMMIE:** *Ja.* I would say maybe about an hour depending how ...

**ADV TEMBEKA NGCUKAITOBI:** Alright, okay, then it is appropriate to adjourn now and then to continue at five to two.

10 **DR ZAID KIMMIE:** Yes.

**ADV TEMBEKA NGCUKAITOBI:** Alright. Thank you. The panel will adjourn until five to two.

**INQUIRY ADJOURNS**

**INQUIRY RESUMES**

**ADV TEMBEKA NGCUKAITOBI:** Good afternoon. Dr Kimmie, we are continuing with your evidence. You were, I think, the current slide, you hadn't started with that slide.

**DR ZAID KIMMIE:** No.

**ADV TEMBEKA NGCUKAITOBI:** *Ja,* you may proceed.

20 **DR ZAID KIMMIE:** Okay. Alright, so many just as a reminder, we're dealing with this particular table which is that global evidence of uneven racial outcome with respect to fraud, waste and abuse and I'm saying that there are a couple of caveats and questions that we need to ask about this sort of result.

First, I think it's important to clarify what is actually meant

when we say that there is a racial bias and then we have to ask whether the results that you've seen are not in fact evidence of racial bias but are an artefact of the classification scheme that I used. So if there were slight change to the classification scheme or even significant changes to the classification scheme would that result that you saw disappear?

So is it just an artefact, something that pops out just because of what we've done and is in fact not really there, it's a constructed result?

10 And then the third question is how reproducible or verifiable is this result? So I want to deal with each of those questions in turn.

When I say that we have evidence of a racial bias, I mean, in the first instance that the bias is meaningful with respect to the racial assignment scheme and I'm using meaningful there as it's a real effect and it's measurable but it is with respect to the racial assignment scheme that we have implemented, so the result is about black versus not black where black has that specific meaning that we've assigned to it based on those list of names.

20 The second part of it means that we have found, as I think we discussed a little bit before lunch, that the racial bias represents a correlation between that classifier in FWA status and we are not making any claims here with this set of data about a causal link or mechanism, we are simply saying that that relationship exists.

And as we said it may be that there's some other factors that explain it but that those factors have to be produced. The assumption

would be upon seeing the result that there is an effect that has to be explained.

And we can infer that this bias in some way underestimates the true level of bias with respect to – if we can think of it, true racial classification. So if you had a gold standard where everyone self-classified, we'll say that because of the way that we've constructed the race variable we would underestimate the true effect of the racial bias and that we can then infer once we've found a meaningful result with respect to this classification that in the real world it would be at least  
10 this and potentially more. So this is the sort of baseline level of deviance in that table.

Then we can ask the question, as I said, is this just an artefact? We can answer it in two ways. Remember that we had a clean racial classifier, that set of names that we harvested from existing sources and it's clean in the sense that that classification has never open the PCNS data file, so it looks only at the list of names and then says make the assignment based on this list of names. So it's a far more restrictive set and under that scheme only about 16% of providers are classified black.

20 If we use that scheme we find a risk ratio – now I'm going to quote those three figures, we don't need the table anymore, the risk ratio of 1.38, the P value of 2, the pie E to the minus 83 and a population risk of 25%. So we would then conclude that there is a very marginal difference between this table and the one that we used with our broader racial classification. If we used the restricted version that

we would still find a deviation of about 40% and that it would have similar impossibility that it would have occurred by chance and so it would be meaningful in the same way.

I won't repeat this exercise for subsequent tables but I have done them for every table. I have rerun the table with a more restricted version and verified that the result is still meaningful by the measures that we have set, right? There is a measurable deviation, the probability is extremely low and the population risk is meaningful and those three factors, they repeat if we use what we can think of as race  
10 1, a restricted version. So there ...[intervenes]

**ADV KERRY WILLIAMS:** Can you explain the restricted version? Ja, okay, define the restricted racial classification.

**DR ZAID KIMMIE:** So the restricted racial classification says we never look at the names in our data set at all, right? The names in that PCNS dataset, they may as well be written in Cyrillic. We only look at names that we can, of external sources, classify black or white or not black and we apply that classification to the dataset so we never look at a name, we never look at Smith and say Smith, you are white and Vesi, you are black. We never look at that, we just look at a list of names  
20 and we make it independent assessment, Smith is white, Vesi is black, etcetera. Then you use that racial classification. So it's far more restrictive because we don't have that huge database of names but the effect size is the same whether we use that restricted version or our expanded version.

The second method would be to say what if we had made

errors in our classification scheme? Now I've – my best guess is that we have about a one percent misclassification rate from not black to black, so there's only in any hundred names only one of those is incorrectly classified as black when it should in fact be not black.

I'm going to assume that we had made a 5% error and in fact that we'd made a 5% error both ways so contrary to what I would believe the case is, I'm going to take a worst case scenario and slightly worse case scenario and say 5% each way. I'm then going to randomly reclassify names and I'm going to do that a thousand times or ten  
10 thousand times and then I'm going to look at the result with that random classification and what we find is that we end up with a risk ratio of 1.36 and a P value of  $2E$  to a minus 101. So that means even if we made an error and a fairly substantial error in our classification, a 5% error, that we still end up with a significant result. We can do this exercise over and over of course. Let's say we had a 15%. What that means is, if I printed out a list of black names in my database, a 100 black names, you would find 15 names that you would say were not black among them, by consensus, and if I printed out a list of a 100 white, not black names, you would find 15 black names among them.  
20 With that sort of misclassification, which is quite egregious, you know, it would be a sloppy system that produced that sort of result, we still have an average risk ratio of 1.26 and a P value that is still infinitesimal.

So I would conclude from this, in fact I checked that we'd have to do a 40% misclassification rate which is basically chance, you'd

have to classify black and not black by tossing a slightly lopsided coin. At that level we can remove the effect. Okay? So we'd have to do something quite odd. So we can conclude from that that the result is robust, it is not due to measurement error with regard to our racial classification.

Okay, can the results be reproduced? So I'm going to treat this by analogy, we're going to make – bake a cake. Let's just say for argument's sake that we're going to bake a carrot cake. I may be partial to carrot cake. I'm going to give you the recipe to make a carrot  
10 cake and tell you what the ingredients are, going to tell you how to put them together, I'm going to tell you how long to bake them, tell you how to mix them and I expect when I do it, I get a delicious carrot cake. The cake may not be all that palatable if your taste don't run to carrots, but it's a carrot cake, we can all agree on that. I'm then going to send you away to bake your own carrot cake. You may not follow the recipe precisely, you may decide that you need more salt, that you need less sugar, that you need gluten-free flour, whatever, but you will end up with carrot cake, you will not end up with a cup of soup when you open the oven, right?

20 So the results can be reproduced. You can take them away and do them yourself, you can reclassify the names, you would not, I think, deviate from the classification we have by more than 5%. You can run the same analysis, the detailed step, every step, document it, you can make slightly different decisions on the margins but even with all of those variations you are not going to account for the scale of

deviation that we observe in the outcome. So this is, I would argue, a real and reproducible effect and as the panel pointed out at the start and the opening statement, what I will make available is that recipe, the detailed steps to construct the datasets and conduct the analysis, the input files, the ingredients, that classification of black names and you're then free to apply it to your own datasets and see whether or not to torture the analogy, you get carrot cake. My bet is that you will.

Okay, now that being said, now the tables get a little bit more messy, unless there are questions on that? Okay, this is the table by  
10 year. So we will construct the datasets in exactly the same way. The universe for the year is all providers were paid during that year, the FWA cases are all cases that were identified as FWA in that year by at least one provider, by at least one of the parties involved, and the various rates are calculate there. So that's enough information to reconstruct the table but again, as we've decided, we'll look at three values. We look at the risk rate per year to give us a sense of scale – so the first thing you'll notice is that the risk rate is lower than the 25% that we had in the previous table and that's because this – every risk rate is per time period, by definition. Risk rate is a time based effect,  
20 what is your risk of contracting something over a time period. So that 25% risk that we saw in the first table, at the population level, was the risk of being identified as FWA over the seven and a half period for which we have data. In this table the risk rate is lower because it's the risk rate per year and you would expect that you would multiply that somewhere approximating the number of years to get the population

risk rate except, of course, that if you appeared in multiple years you're only going to count once so, you know, that's where all of it sort of comes out in the wash.

**ADV KERRY WILLIAMS:** For 2019 it's just for six months.

**DR ZAID KIMMIE:** Yes, so the 2019 figures are lower because they're for six months. What is not a time based variable is the risk rate. Right? The risk rate is impervious to the time period. That's because by simple algebra it appears in the numerator and the denominator. So when you divide those two quantities, the time period disappears. The  
10 risk rate is always the risk rate, it doesn't matter whether it's calculated over ten years or six months and we can see there that the risk rate has increased – it started off quite low and you'd argue potentially that that 2012 figure, that P value is low, but it's not – you know ...[intervenes]

**ADV KERRY WILLIAMS:** Talking about the ratio, risk ratio?

**DR ZAID KIMMIE:** Yes, the risk ratio, 1.22 in 2012.

**ADV KERRY WILLIAMS:** It goes up to almost twice.

**DR ZAID KIMMIE:** Yes.

**ADV KERRY WILLIAMS:** By 2017.

20 **DR ZAID KIMMIE:** 2017 and then it declines quite markedly in 2018 and the data available for 2019 indicate that there is in fact no difference. If we only looked at 2019 data, we have a risk ratio of 1.1, which is small, relatively small, we'd have a P value of 0.04, which means it's not significant, so we would dismiss that as a meaningful result. So in 2019 we'd say nothing is going on. And it's still, even



though it's six months, it's still a result. But in every year basically, from 2012 through to 2018 we have a deviation and the risk ratio at very significant levels.

**ADV KERRY WILLIAMS:** Do you have an understanding of what's happening in 2019 or why this is happening in 2019?

**DR ZAID KIMMIE:** No. I have no input. I mean, my sense is that perhaps sort of offhand remarks that having the panel of inquiry has somehow put the brakes on the fraud identification process as a whole. So there's potentially that explanation but, I mean, that would just be  
10 speculation.

**ADV KERRY WILLIAMS:** The same for 2017 you wouldn't know would account for that...

**DR ZAID KIMMIE:** Yes.

**ADV KERRY WILLIAMS:** ....very increased risk ratio.

**DR ZAID KIMMIE:** I would say one more thing here. That last point, it's just a technical point, but just in case someone were wondering, when you do tests of this form you do a basket of tests, you adjust the P values to take account of the fact that you do multiple tests. So there are standardised ways of doing that adjustment and these values  
20 have been adjusted to take account of multiple testing. So it's basically when you do multiple testing it's like taking multiple bites at the cherry and you're more likely to pick out something just by chance. So there's a – we've taken care of the fact that we do multiple tests in this table and every table where that sort of thing occurs. Okay.

What the conclusion would be from this table would be that the

effect that we saw in that global table is not due to an adverse outcome in a single year, it's consistent over time and in fact the pattern increased as time passed from really what would be a relatively low 22% in 2012 up to almost double the risk rate in 2017. Okay.

This table is the same data cut by discipline, right? So in order to read this again, you would simply read the risk ratio and the P value and then as a sanity check what the population risk rate is. So just for GPs, if we read that, over this seven and a half year period, about a quarter of GPs had been identified as fraud, as FAW cases.

10 That difference between black and not black is now in the order of 13% and that is a risk ratio of 1.6. Remember that the population risk ratio was 1.4, so this is more extreme than what was found in the population and again it is a significant result.

The second line there's pharmacies and what you would notice about pharmacies almost immediately is that there are 4 500 of them and 600 pharmacies are classified as black. And this is a result of our classification scheme, almost 80% of pharmacies did not have the details necessary to classify them, so they were white. So this is a case where that classification, if there had been an effect, has wiped it  
20 out, that in fact black pharmacists are identified at a lower rate than not black pharmacists, 44% compared to 52%. The rate at which pharmacies are identified, though, globally, is very high. Over the seven year period more than half of all the pharmacists identified was FWA. And again there's a calculation of a risk rate at a P value.

So reading this sort of table gives you a sense of where the

deviations are highest by discipline. Social worker, for example, third from the bottom, those are relatively small numbers, 1 500 people, but the risk ratio is 4.5. 33% of black social workers were identified as FWA cases compared to 7.5% of not black social workers and so on and so on.

So again this is where it becomes a bit of a rabbit hole, you can carry on doing endless tables of this sort but this is a useful indication of the differential rate at which particular disciplines are being identified as FWA and within that, the different application  
10 relative to whether you are black or not black.

**ADV TEMBEKA NGCUKAITOBI:** You ended with a pharmacy and then you were dealing with social worker but you haven't gone through optometrists and physiotherapists.

**DR ZAID KIMMIE:** Yes, again, it's simply a matter of saying optometrists, that's 1.4 – and that's pretty much what the population risk ratio is, so that's no big deviation from the population finding. There is a sense there that perhaps – I think about a quarter of optometrists were not able to be classified, so in optometrists as well we have a problem where the misclassification may have obscured any  
20 effect that existed. Physiotherapists is also – that's quite a severe ...[intervenes]

**ADV TEMBEKA NGCUKAITOBI:** Ja, but I mean that's only the risk ratio, if you look at the absolute numbers on risk it's still 28.7% versus 20.5%.

**DR ZAID KIMMIE:** Yes.

**ADV TEMBEKA NGCUKAITOBI:** And that still accounts for approximately 7%.

**DR ZAID KIMMIE:** Ja, so that risk ratio of 1.4 is the same as the population risk ratio. So we'd say even with those caveats, it would appear that the situation among optometrists mirrors that of the population as a whole with that 40% increase in risk.

The physiotherapists has a risk ratio of 2.1 so black physiotherapists are targeted at twice the rate of not black physiotherapist, it's double the risk rate, give you a risk ratio of 2 and  
10 you can see that again in the scale of the figures, amongst all physiotherapists, about 19% were identified as FWA cases. The differential between black and not black is 32 to 15, double the rate.

Dentists, there's a case where there is no deviation. The risk ratio is 1, so it's what you'd expect if there were no effect of the racial classification at all and the P value is 1, it's a completely – so we would accept in that case, absent any further information, that the null holds, that there is no racial bias.

**ADV KERRY WILLIAMS:** What's your – do you have an explanation for that?

20 **DR ZAID KIMMIE:** No, I mean, at this point it would need some investigation of what the actual practices were in that specific discipline. And remember, I'm putting up in these tables the list of disciplines that have reasonable counts. So, I mean, if you've looked at the classification scheme for disciplines, it runs to three or four pages, you know, you can get ambulance, emergency ambulance,

etcetera, etcetera, so it's the – I'm only picking out for this table the large numbers of cases.

Independent specialists, again there seems to be no effect at all. What small effect there are is in fact the converse, non black specialists are targeted at – no, sorry, the effect is quite ...[intervenes]

**ADV TEMBEKA NGCUKAITOBI:** It's still at 46, yes.

**DR ZAID KIMMIE:** Yes.

**ADV TEMBEKA NGCUKAITOBI:** 46 versus 40.

**DR ZAID KIMMIE:** 46 versus 40 but this is a case where you'd say the  
10 risk rate is quite small, it's that 16%, and the P value is really quite high, so you would dismiss this as something that potentially could have occurred simply by chance, not necessarily find anything meaningful in that deviation.

**ADV KERRY WILLIAMS:** So, in other words, it could have happened by chance.

**DR ZAID KIMMIE:** Yes.

**ADV KERRY WILLIAMS:** It's not meaningful, it's not meaningful.

**DR ZAID KIMMIE:** Yes.

**ADV KERRY WILLIAMS:** Correct.

20 **DR ZAID KIMMIE:** Ja, if you had to make the conclusion here that amongst independent specialists there is no evidence within the data of racial bias. Okay.

**ADV KERRY WILLIAMS:** You're not leaving the slide, though, are you?

**DR ZAID KIMMIE:** No, no.

**ADV KERRY WILLIAMS:** Okay.

**DR ZAID KIMMIE**: Okay. Again I would guess that you'd have more detailed information from the testimonies about experience of particular specialists. Psychologist is also an important case. I think after GPs psychologists are I think the second largest group of practitioners in the classification scheme. The deviation there is extreme. It's at three and a third, close to three and a third and you can see that 8% amongst white psychologists, 26% amongst black psychologists. That needs, as the saying goes, the more – the higher the level of deviation, the greater the explanation that would be required to justify it.

10 **ADV KERRY WILLIAMS**: In other words here black psychologists are more than three times likely to be investigated.

**DR ZAID KIMMIE**: Yes. Again, then you use as your guide there that 2E to the minus 60, which, you know, again is in puzzle putting itself together territory, just not – this didn't happen by chance.

Anaesthetist, you know, I'd be taxing my medical knowledge slightly, but anaesthetists usually work as part of a team, some surgeon or other conducting – I think it's a more restricted sort of a practice.

20 Again, obstetrics, it's a high risk ratio, 1.45 but P value is now out of the impossible range and now just in the extremely unlikely range, you know, 1 E to the minus 6 is still an extremely small number but the principle we're going to follow is if you're going to judgment you'll need like strong evidence. That would still count but it would be more of a borderline case. We're not – we don't have a cut-off for P value, right? It's always a combination of things.

Social workers is the one I mentioned earlier that's – the last

three in fact are all instances where they are very high risk ratios, double that of the population and that means that if you read them as percentages, that stark difference between black and nonblack. 33 compared to 7.5 for social workers. 50% of registered counsellors. Those are small numbers, you know, there are only 857 registered counsellors but of the black registered counsellors, just about half of them have been identified as FWA cases. That's a pretty severe and there are probably, I would guess, medical explanations or clinical explanations for why some of these population risk rates are so high.

10           Amongst obstetricians for example, the fact that 50% of obstetricians have been identified as FWA cases, I would assume that there is some clinical explanation for why that has happened rather than...

**ADV ADILA HASSIM:** [Indistinct]

**DR ZAID KIMMIE:** It is just a very high rate. One in two chance that you have been identified as FWA in a practice, I mean, that is pretty high, you know. You can flip a coin whether you have, I mean, you know. By a simple probability. Again, there may be other explanations, not necessarily for the deviation but for the population risk rate.

20           Perhaps there are practices that people have identified.

**ADV TEMBEKA NGCUKAITOBI:** You would expect there that there should be some correlation between the Black versus White and the population risk ratio. The problem is the deviation between the population risk ratio and the Black risk ratio.

**DR ZAID KIMMIE:** Yes – is extreme, of the same order of magnitude as the population deviation at 1.4 that we identified. Alright, again I think, *ja*, there is a lot of information in the table. The conclusion that you would want to draw from this is that, where that deviation has occurred it is not located within a single discipline or a small group of disciplines. The phenomenon we are looking at is widespread. It is not as if, for example, it was only GP's which had this differential rate and everything else was fine. And the GP's were responsible for the population deviation. So, we have really had ... *ja*?

10 **ADV KERRY WILLIAMS:** Except for the fact that it seems that specialists are out of the picture.

**DR ZAID KIMMIE:** Yes. Yes.

**ADV KERRY WILLIAMS:** Yes and that is a large number of sub-specialities which sit in there.

**DR ZAID KIMMIE:** Yes. So we are not saying that it happens in every discipline but you are saying it is widespread enough that you can dismiss the idea that that population effect is simply due to one subgroup or subset of disciplines that result in the population deviation, that it happens across a wide set of disciplines. Okay.

20 This is slightly more contentious then. We can do exactly the same sort of analysis but restrict the analysis to each of the parties that contributed data. And I know I have asked for guidance and you have indicated that there is information in this case that cannot be shared generally, that is potentially still confidential. But the same principle applies, the population that we are looking at for each of the



administrators is the total set of providers that were paid during that period 2012 to 2019. And the FWA cases of the total number of unique cases identified during that period. So if a provider identified, if a provider appears more than once in this period, we will count it as a single case, so we are not going to double count a provider because they have occurred in multiple years. It will count as a single case so we are looking at a set of providers and a subset who have been identified as FWA cases.

So those are the risk ratios. So what I have removed from  
10 there is all of the numbers that allow you to calculate the individual risk rate. And all we have ...

**ADV ADILA HASSIM:** Just to be clear, it is the middle column in the first block?

**DR ZAID KIMMIE:** Yes, the number of FWA cases over the period - that is removed. And removing leaves you at a slight disadvantage, well not you because you can see the table, but from the audience perspective because you do not know what the population risk rate is and by the sort of guidance that I have laid out you should look at three numbers to make the judgement. The population risk rate, the risk ratio, and p-  
20 value. But here we can assume that we know that the risk rates are meaningful and that the risk ratios then across the three providers in each case is sufficiently high to indicate that it is meaningful and sufficiently unlikely to say that we can eliminate chance as an explanation.

There is an effect across all of the providers although the scale of that effect differs according to each of those providers. And that it ranges from 1.35 for Discovery Health up to 3.31 with those p-values. I think that is the smallest p-value I have seen in this particular set of analysis.

**ADV ADILA HASSIM:** Sorry, which is?

**DR ZAID KIMMIE:** E to the minus 205. *Ja.* That is, cut someone in half in a magic show ....

**ADV ADILA HASSIM:** That is outside of jigsaw puzzle territory?

10 **DR ZAID KIMMIE:** Yes. *Ja.* You know it is a little bit of overkill at this point because unlike, you know, we are talking about events that are unlikely. If it is E to the 100 or E to the minus 200 we are splitting hairs, it is simply not possible.

**ADV TEMBEKA NGCUKAITOBI:** Well I suppose it does add into the full picture because what we now know from your exercise is that we know per discipline that the pattern of racial bias is exhibited. We also know that each and every year that pattern is repeated but we now know that across the schemes that you looked at, those patterns are replicated.

20 **DR ZAID KIMMIE:** Yes. And that would be that table so that is by provider, by year. Again the information that would allow you to do the calculation on, you cannot reconstruct that two by two table using this information. So it is just the statistics attached to the table that pop out with the exception of the population risk.

**ADV TEMBEKA NGCUKAITOBI:** What is the reason why the risk ratio of Medscheme only starts in 2016?

**DR ZAID KIMMIE:** The numbers are too small to make for meaningful calculations.

**ADV TEMBEKA NGCUKAITOBI:** Numbers prior to 2016?

**DR ZAID KIMMIE:** Yes. That is a combination of incomplete data, so the explanation that was provided was that the data could not be reconstructed for that period. So they had information on some FWA cases but certainly not all. And so, I have removed that because the p-values and risk ratios would be misleading in that.

**ADV ADILA HASSIM:** Does that explain the 2017, 1.82 in the earlier  
10 table? Could that, is it because the data from Medscheme, I mean do these two ...

**DR ZAID KIMMIE:** No.

**ADV ADILA HASSIM:** ... not work together at all?

**DR ZAID KIMMIE:** *Ja.*

**ADV ADILA HASSIM:** Do you know what I mean, the 2017, 1.82 risk ratio on slide 39, where the risk ratio went up to almost 2 in 2017?

**DR ZAID KIMMIE:** Yes you could infer that, yes, that Medscheme played a greater role in driving up that population risk ratio because their risk ratios are above what was observed for the population ...

20 **ADV ADILA HASSIM:** Yes.

**DR ZAID KIMMIE:** ... in those years.

**ADV TEMBEKA NGCUKAITOBI:** I wonder, if you look at the Discovery table what you see in 2012, I mean no, I suppose the complaint there is 1.09, that is the risk ratio. 2013 -1.12, 2014 is now on the questionable level and so is 2015 and 2016. 2017 seems to be out of kilter but 2018

then goes back. But 2019 – you said we can exclude that. So although earlier I suggest that across the schemes these facts are replicated over the years, it does appear that the real problem is between the years 2014 to 2018. I suppose that is still a long period. That is five years.

**DR ZAID KIMMIE:** *Ja.* Just to correct that, the risk ratio over the 29 period is still meaningful because it is independent of time. The risk rates would have to be adjusted but that risk ratio is still a real measure. So, the fact that it is 0.9 in 2019 with a p-value of 0.07  
10 means the effect that was visible in earlier years had dissipated in the 2019 data. There is no racial bias evident in that 29 data as it stands.

**ADV ADILA HASSIM:** In general the risk ratio on the part of Discovery is much lower than the others.

**DR ZAID KIMMIE:** Yes. Significantly lower.

**ADV KERRY WILLIAMS:** Then again if you look at the 29 p-value and risk ratio for Medscheme for 2019 it hasn't got any lower.

**DR ZAID KIMMIE:** No.

**ADV TEMBEKA NGCUKAITOBI:** In fact the Medscheme risk ratio has gone up from 2018 and it stayed the same for GEMS.

20 **DR ZAID KIMMIE:** Yes.

**ADV TEMBEKA NGCUKAITOBI:** It might show that you know your earlier speculation that this inquiry has had an impact actually is neutral.

**DR ZAID KIMMIE:** I am not sure I understand?

**ADV TEMBEKA NGCUKAITOBI:** When you earlier speculated that the reasons we are seeing lower rates of Black FWA being targeted is perhaps the existence of this inquiry. That this inquiry, in fact, has had a neutral effect on what the schemes are doing.

**DR ZAID KIMMIE:** Yes. You say the effect may be uneven.

**ADV TEMBEKA NGCUKAITOBI:** *Ja*, I mean you can say for Discovery it is obvious that it has been a positive effect but for the other two it has been neutral.

**DR ZAID KIMMIE:** Yes, and there may be explanations linked to the  
10 precise way in which the systems are implemented. That would be worth talking about. *Ja*. I don't know if there is much else to say about this particular table. Okay. What does appear to have disappeared for some reason is the table by provider and administrator, by administrator and discipline. But that I think perhaps is a little bit too detailed in any case for easy consumption. Okay.

So just to recap on the issues that were raised; there is no evidence of explicit racial profiling in the design or implementation of systems used to identify potential FWA case by the ...[Intervenes]

**ADV ADILA HASSIM:** Dr Kimmie on that table of administrator and  
20 discipline was there nothing that you wanted to highlight in relation to that table?

**DR ZAID KIMMIE:** I would say a gentle comment is that it does reveal the very specific implementation of FWA systems across the various disciplines and to give a general example, if, whatever system people were implementing was modular in the sense that you could apply that

fraud identification system as an analysis tool only to specific disciplines then it would demonstrate that the chances of identifying those disciplines would be higher because you would have an additional tool. So for example, if ...

**ADV KERRY WILLIAMS:** Could you just start by explaining what you mean by modular.

**DR ZAID KIMMIE:** Okay, so if administrator A was implementing a fraud detection system and decided that they would either buy or build the GP module first and then they would do a pharmacist model and  
10 then they would move on to the others as they became, as the cost justified the additional investment in modules. And that would mean that if all you had for example, was a module that looked at GP's and pharmacists, that you would target GP's and pharmacists because that is what your data would be guiding you towards. That if you had no module looking at dentists or optometrists then you would have no information about those areas.

And in fact, it is the case that each of the parties has implemented different parts of their fraud detection system at different times. It is, none of the three has a complete solution dropped in.  
20 They have all been built over time. Some of the patterns reflected in the way the particular parties have focussed their efforts on particular disciplines reflect that set of choices.

**ADV TEMBEKA NGCUKAITOBI:** We have got a copy of this table, it is table 5.6. Do want to put it up and talk through it so that, because you say it has disappeared.

**DR ZAID KIMMIE:** I not sure that it was, all I have is the version that we can't share.

**ADV TEMBEKA NGCUKAITOBI:** No, this version seems to be fine.

**DR ZAID KIMMIE:** No.

**ADV TEMBEKA NGCUKAITOBI:** Alright, well so which one do you want to deal with? Do you have the version that you had prepared for your powerpoint?

**DR ZAID KIMMIE:** No, it seems to have, that is the, it would take me a bit of time to find that.

10 **ADV TEMBEKA NGCUKAITOBI:** To reconstruct it?

**DR ZAID KIMMIE:** *Ja.*

**ADV TEMBEKA NGCUKAITOBI:** Alright, well anyway, if you have got this you can look at this and then tell us what conclusions we should draw from this.

**DR ZAID KIMMIE:** Okay.

**ADV ADILA HASSIM:** I would like you to do that as well because there are some very interesting conclusions I think to be drawn in relation to which disciplines, risk ratio in relation to discipline.

**DR ZAID KIMMIE:** Okay, so I mean ...

20 **ADV ADILA HASSIM:** And each administrator's role in that discipline.

**DR ZAID KIMMIE:** Yes. Okay, so I am going to read off some of these figures just for interest sake. So within GP's, the general practitioner, which is the largest group, the risk ratio is above 1.4 for all three of the parties, Discovery, GEMS and Medscheme but it is significantly higher for GEMS and Medscheme. So it is 1.37 for Discovery but rising up to

2.2 for GEMS and then 2.3 for Medscheme. So the deviation is more extreme for GEMS and Medscheme for GP's but all three exhibit extreme deviation. I think an interesting observation with pharmacies, again we mentioned the classification problem and that at the population level it would appear that White, not-Black pharmacies are being targeted at a higher rate than Black pharmacies. That pattern is in fact reversed for Medscheme and even with that restricted classification, Medscheme, Black pharmacies in Medscheme have a three times higher rate, risk of being identified as FWA than non-Black pharmacies within Medscheme.

And that, ...yes?

**ADV ADILA HASSIM:** I am waiting to hear what you have to say about physiotherapists.

**DR ZAID KIMMIE:** [laughs] Amongst optometrists is again, and this is an indication if you will, of priorities and investment, GEMS has again a risk ratio of 2 for optometrists and for both other providers the risk ratio is to all intents and purposes – 1. The actual risk ratio for GEMS is 4 but the p-value makes the actual count meaningless.

**ADV TEMBEKA NGCUKAITOBI:** Who are you talking about because GEMS is 1.98 and then Medscheme is then 3.5.

**DR ZAID KIMMIE:** Yes, sorry, *ja*. So, Medscheme for optometrists has a very high risk rate but it is meaningless. It should be ignored. GEMS has a risk ratio of 2 which is meaningful, that is the differential rate at which Black optometrists are being targeted and again for Discovery, we can ignore the risk ratio. It is to all intents and purposes – 1.



Physiotherapists is an extreme example, again as I said, physiotherapists consists of I think the third largest group of practitioners after GP's and psychologists. The risk ratio for Discovery is 2 so that is quite severe. For GEMS it is 6 and for Medschemes it is 13. *Ja*, that, *ja*, is a very serious deviation. Dentists again is a case where GEMS is a risk ratio that is meaningful of 2. It is pretty much 1 for Discovery and Medscheme. Independent specialists, Discovery is again 1 so it is flat and the deviation occurs with GEMS and Medscheme with both, both have risk ratios of about 2.

10 Psychologists we mentioned before, the second largest category, the risk ratios are all of the same ballpark. For Discovery and GEMS at about 4 and Medschemes – 6. So remember what risk ratios of 4 and 6 mean. Four hundred times more likely, four hundred percent more likely, six hundred percent more likely to be targeted. We are talking about differences of 6 for non-Black compared to 25 for Black.

**ADV TEMBEKA NGCUKAITOBI:** Physiotherapists, does it mean they are 1200% more likely to be targeted by Medscheme?

**DR ZAID KIMMIE:** Yes, the twelve.

20 **ADV TEMBEKA NGCUKAITOBI:** I was wondering what the twelve means, it is one thousand two hundred?

**DR ZAID KIMMIE:** Yes. Twelve times. Anaesthetist is ...

**ADV ADILA HASSIM:** One thousand two hundred times more likely?

**DR ZAID KIMMIE:** Which number are you talking about? Twelve times more likely. I think that is the best way to read it. Anaesthetist there is

nothing happening. The obstetrics, there is an effect, a very small effect amongst Discovery 1.3, 1.4 but potentially we can ignore that. Amongst GEMS it is 4, a risk ratio of four and Medscheme we can ignore because it is not a meaningful measure. I think those last three are again cases where we saw the population risk ratios being extreme – amongst social workers it was somewhere around 7%, I mean seven times, a risk ratio of seven and that is replicated in the data. All three, across all three providers.

**ADV TEMBEKA NGCUKAITOBI:** It is not 7%, it is 700%.

10 **DR ZAID KIMMIE:** Yes, the risk ratio is 7. And that pattern happens, there is something, I would say there is something material happening with social workers. Registered counsellors, again they are a slightly smaller group. They are only about 600, between 600 and 690 registered counsellors. The risk ratio for Discovery is 12. One extremely high.

**ADV ADILA HASSIM:** No, it is 5.

**DR ZAID KIMMIE:** Sorry, I am reading the wrong column, 5 and with GEMS it is 2 and Medschemes, 3. GEMS again you can ignore. So registered counsellors - Discovery and Medscheme. And then dietitian,  
20 there is quite a marked deviation in all three groups. Discovery 5, GEMS 4, Medschemes 10. All very high rates with very low probabilities.

**ADV TEMBEKA NGCUKAITOBI:** Will you prepare the slide so that it is included as part of the [Indistinct]

**DR ZAID KIMMIE:** Yes, okay. So I think there is a lot more data

to analyse but I think this is sufficient to make believe the case that there is clear and strong evidence of a racial bias with respect to outcomes in FWA processes. I think that is not really in any doubt. I believe that that result will hold up irrespective of the marginal changes that may be made to the way the algorithm is implemented, my analysis is implemented. But that will be available for people to inspect in detail. I indicated the bias is not restricted to a limited time period, nor is it located within only particular disciplines. They vary in scale across these factors but it is a pervasive, widespread effect. It is not a  
10 bad apple or a bad set of disciplines or a bad year. This is more than an outlier effect in a limited time period or a limited set of disciplines that have given rise to that global effect.

This is something that is pretty widespread. And I am convinced, because of carefully examined assumptions that underpin these findings, and I am convinced these findings are sufficiently robust, particularly given the cautious way in which we have proceeded with the classification. That there is sufficient that is indicative of a real phenomenon that needs explanation and that any reasonable classification scheme it will give rise to similar results.

20

What would be an unreasonable classification scheme? I did one, I used the fact whether the last digit in the PCNS number was odd or even and I classified people, the providers into groups called red and blue, based on whether the last number in the code was odd or even. That you would say is blind to any demographic and as you

would expect, when you do the analysis there is nothing to be found, nothing. The p-values are close to one, the risk ratios are close to one. Any random assignment of race, if race were not a factor, if we randomly assigned people to Black and not-Black you could run that assignment twenty, thirty, a hundred thousand times. You would not get close to the data we have seen here. So again, it is another way of saying this is not an artefact of our classification. This is a phenomenon that is real and needs to be explained. Oh, and of course, the first point I would repeat, there is no evidence that this is  
10 implemented as part of the fraud identification systems. It may be the case that the factors that are used are linked to race, that they tend to identify Black providers more often than not-Black providers but that would need further investigation to identify where exactly that mechanism has occurred.

So when I say there is no explicit racial profiling I am not ruling out that the actual factors, the analytics used may contribute to a racially biased outcome. But that, in the absence of information about race, the parties involved at this stage are unable to consider the question of whether there is a racial bias.

20 **ADV KERRY WILLIAMS:** On that point, there is only one entity that has made their algorithm available as I understand it. The other two, we don't have insight into those algorithms, is that right?

**DR ZAID KIMMIE:** *Ja.* There is one set of algorithms that is detailed enough that, given a racial classification, you can start asking the question about whether the algorithm itself will produce racially biased

outcomes but the other cases are black boxes. So there is no, well, there is one black box and there is one grey box which has a combination of built and in-built factors. But even in that case the precise mechanics of how the result is calculated is proprietary to the service provider. And we have, *ja*, we would be not be able to ask or understand how that happened.

**ADV TEMBEKA NGCUKAITOBI:** Thank you Dr Kimmie. It remains of me then to thank you on behalf of the panel for a thorough job that you have performed and making our job a lot more easier, I hope so,  
10 otherwise than it would have been. Obviously you still remain available to the panel for questions to be asked and further enquiries to be conducted. Now the future of this inquiry is that we have now completed the public hearings in relation to the complaints that were made against the schemes. We have also completed the evidence of experts.

The next big phase will be to hear from the schemes. We had suggested preliminary dates but those dates for various reasons couldn't be met. We will be communicating further dates to all the representatives of the schemes and just to emphasise the point we  
20 made at the beginning, we are anxious that the matter should be finalised as soon reasonably practical but we appreciate that people have rights to a fair hearing. They would like to engage with the data that we have and today's evidence is heard by many people for the first time and they want to engage with that as well. But despite all of those shortcomings we would appreciate the people making themselves

available on dates that will be communicated by the secretariat. We are adjourned until further notice.

**INQUIRY ADJOURNS *SINE DIE***

**TRANSCRIBERS CERTIFICATE FOR**  
**THE COUNCIL FOR MEDICAL SCHEMES (CMS) INQUIRY UNDER**  
**SECTION 59 OF THE MEDICAL SCHEMES ACT**  
**HELD AT**  
**BLOCK A, ECO GLADES 2, OFFICE PARK, CENTURION**

DATE HELD : 2019-11-19

10 DAY: : 14

TRANSCRIBERS : V FAASEN; B DODD; C LEHMANN

**Audio's are typed verbatim, as far as audible/possible**



# Accura Africa Group

Your Forward-Thinking Partner