Policy Pitch - *Safety and Quality – time to stop just goggling at the problems our data reveals*

Melbourne 15 August 2017

The health system is awash with data. Nearly all of it tells us that we could improve. It usually also tells us where we need to improve; particular places and patients and care processes. Yet our current improvement efforts are surprisingly limited in their scope. One major reason is the lack of actionability of many kinds of health data. Soundness, relevance, accessibility and understanding of health data can all be improved. In this Policy Pitch event, a panel of experts discussed this improvement process.

Speakers: Stephen Duckett, Director of the Grattan Institute Health Program
Christine Jorm, Grattan Health Program Advisory Group

STEPHEN DUCKETT: I'd like to acknowledge the traditional owners of the land, the Wurundjeri people of the Kulin nation, and pay my respects to elders past and present, as well as emerging elders of the nation. I'd also like to remind us that Aboriginal people have a life expectancy of ten years shorter than non-Indigenous Australians and that when we hear those acknowledgements of country we should remember that they're not acknowledgments of rote, but acknowledgments to remind us of the continuing dispossession and the problems associated with that.

One of the foci of the Grattan Health Program is quality of care and we've done one report which looked at one aspect of that, which is what I'll talk about tonight, and we've got a couple of other reports underway, so what Christine will be talking about is some of the work that we're doing right now to lead to these reports which we hope to release later this year. You're a test run, so thank you very much for being unwitting volunteers and guineapigs, and we'd really appreciate later to hear your reflections on what we have to say and any comments that would be helpful for us as we keep on working on the reports. My presentation will be on a report we issued last year on so-called questionable care and what we're interested in is the problem that there's been so much published about this and so little action. One of the themes of our forthcoming reports will be this issue of actionable data. I'll start with this presentation. This is work done by a guy called Jeff Richardson, who was Professor of Health Economics at Monash University. About 20 years ago he published this graph which shows the admission rate for various operations across Victoria, these are local government areas and you can see huge variations in admission rates. This issue of variability is a really common and high profile one. Just a month or so ago the Australian Commission on Safety & Quality in Healthcare issued yet another report called *The Atlas of Variation* which did exactly this. It's in a somewhat prettier form, but it's exactly this, publishing information about variation. So we should be celebrating the 20th birthday of this publication and also, of course, the 20th birthday of nothing happening as a result of that publication.

Now, this reminds me of the very famous Sherlock Holmes story *The Dog in the Night-Time* where we have the police inspector saying, "Is there any other point to which you wish to draw my attention?" and Sherlock Holmes says, "To the curious incident of the dog in the night-time". Gregory, "The dog did nothing in the night-time". "That was the curious incident." So what we have to explain is why something didn't happen. Why I put that up is Bob Evans, who is a very famous Canadian health economist, probably *the* most famous Canadian health economist, published an article in 1990 where he says, amongst other things, "Knowing is not the same as doing. The most striking fact about the large and

extensively documented variations in pattern of medical practice throughout the developed world is the minimal impact that this information has had on health policy". Then he said, "These differences, moreover, do not appear to be explicable in terms of the needs or characteristics of the population served, at least such explanations have not been found". So he's saying the same thing. He then goes on to say, "Similar observations emerge in every country studied, despite the substantial differences in their health systems. The research findings are really quite remarkable in their scale and consistency, yet nowhere has there been any significant adaption in regulation or reimbursement" and, back in 1990, he also said, "But now it is 20 years since the first studies in the English-speaking world began to focus on this phenomenon" quoting Pearson in '68. So now we've got the 50th birthday of this phenomenon being published and the 50th birthday, of course, of nothing happening about it.

In that tradition, we decided to do exactly the same thing, so we, of course, decided to waste our time doing a graph which shows exactly the same thing, huge variation. This time the dots here are what are called Medicare Local areas and there are 61 of those across the country. Huge variation in tonsillectomy rates, for example, from one that had a rate which was more than double the national average to one which was 75% below the national average. So huge variations we're observing. The puzzle is - well, it's probably not a puzzle, but one of the issues is why doesn't anything happen? One of the problems is we've got no idea whether this is good or that is good. So we can array ourselves, but high could be bad and low could be bad and we don't know is it due to differences in cases, is it due to difference in patient preferences or is it due to severity variation? So the problem is it doesn't tell you anything and it starts to ask questions, but then where do you go with those questions? There's no clarity about what it shows and so the question is why on earth do we keep doing it and even if you knew what variation was legitimate, what could you do about it? So what we said was what happens now?

Typically, these atlases of variation focus on where people live. What we showed you was geographic variation. What Jeff Richardson published was geographic variation. What Bob Evans talked about was geographic variation. That's the conventional analysis, that's what the *Atlas of Healthcare Variation* publishes, geographic variation. One of the beauties of that is you can admire the problem. Ours is pretty with orange. There are others which do it beautifully with maps and you have different colours for the different rates and it's beautiful. That's good; I'm not against beautiful things. It has the issue of what is good but, most importantly, there's no accountability mechanism whatsoever. So who is accountable if we discover that Kew has a high rate of admissions for tonsillectomy? What is the mechanism for doing anything about having identified that as an issue and turning it into something to do? In Victoria we have public hospitals not responsible for a particular area. In New South Wales and Queensland, traditionally the organisational mechanism that is responsible for public hospitals have a geographic responsibility. In Queensland they're called local health and hospital districts and they used to be called area health authorities in New South Wales. The very title of these things gave them a geographic orientation and there were boundaries for these organisations. Importantly, there's been no accountability even in New South Wales and Queensland, which have had area health authorities for a long time.

Why is that? Because they don't have responsibility, for example, for private hospitals and some of that admission rate is people being admitted to private hospitals. They have a responsibility for part of the issue. No-one has responsibility for the whole issue. So we said it's crazy to keep on analysing at the area level. As the great saying goes, "insanity means doing the same things over and over again and expecting different results" attributed there to Albert Einstein. In the interests of full disclosure I'll point

out there's no evidence that Albert Einstein ever said that, but it's a lovely quote and the evidence that he didn't ever say it is here. It is a fantastic quote but, of course, not everything you see on the internet is true. So our approach was to go on its side, to say let's look at the hospitals where this happened and once you start looking at the hospitals you can hold someone to account. So what we started to do was to move our analysis away from the geographic area where variation suggests there might be inappropriate care, to begin to look at those conditions where you might have more certainty that variation is an issue. So you might ask can we identify treatments for which, on average, this treatment should not be routinely provided or there's high level evidence that, on average, this treatment should not be provided or definitive advice, so being a bit more certain each time you went up that spectrum.

We identified five conditions for which there was basically no evidence that they'd contributed positively to a patient outcome. This is not to say there's no evidence that they didn't contribute positively on something else. They might have contributed positively to the income of an orthopaedic surgeon or the income of someone else, but there's no evidence that they contributed positively to patient outcome. So there were a number of these done across the country and we also looked not only at things we called "do not dos" but "do not dos routinely" where the advice was you shouldn't do these things on a routine basis. We found that there are a number of hospitals in all states which admit to doing it to at least one of these to people, sometimes they have two types of those, sometimes three and sometimes four or five of them, but what we have here is instead of these dots representing areas, which they did in our first slide, these dots now represent hospitals. If we look at the arthroscopy example, this is arthroscopy of the knee for osteoarthritis and it's pretty clear that there's no evidence that this works. This black bar is the average rate, so each dot here is a hospital which both admits patients for osteoarthritis of the knee and also does arthroscopies. This is the rate across the country and about 2% of all admissions for osteoarthritis of the knee had an arthroscopy, but you see there's one hospital up here where 70% of all their patients admitted for osteoarthritis had an arthroscopy. So a huge variation, bearing in mind the right rate here is zero. We've got rid of the problem of knowing whether a high rate is good or a low rate is good; the right rate is zero.

So what we said is we should have a process where we start looking at these high rates. So the issue that we're concerned about is now looking at variation and identifying where something goes wrong or where there are high rates and start to focus on this. What do you do in the first instance? Well, you tell the hospital, "Your practice is abhorrent. You're doing something differently from your colleagues". They may not know that their practice is different. These are public hospitals, by the way, so we're not looking at financial influences here. So they may not know that their practice is so abhorrent. Then you say that providing information is often not enough, so the second stage should be do they still have a high rate a year or so later, when there's a chance to have actually thought about their practices? So you give them the information back again and then you start to ask what do we do if they still have a high rate? Our recommendation in this report was that you send a team of their peers in to ask why was it in this particular case, this particular patient, you decided to do an arthroscopy. Because it may be that there might have been reasons that we know nothing about for why that decision was justified, but it's all about holding individual clinicians to account for the decisions they made and also holding the hospital to account for the decisions that were made within their hospital.

So the big difference between what our approach is and what the Atlas of Variation and all these other geographic approaches are, there's not an institution that you can hold to account and, of course, the same applies to those other five conditions for which there is little or no evidence. So these are the so-called "do not dos". These ones are the "do not do routinely", so these ones are where the evidence is

not so clear. The evidence is not that you shouldn't do this, but the evidence is this should not be done routinely. Obviously, as soon as you move from a "do not do" to a "do not do routinely" you move to a situation where there's much greater variation and so the average now is much greater than the average of those "do not dos". In the previous case we said that everybody above the average should be looked at. This time we said maybe only the people in the top 10% of the distribution should be looked at, but still there is a mechanism to hold to account the organisation within which these things happen. So the big message of this part of the presentation is this: what we've done here is transformed the data from being ornamental, something to look at, into something that is actionable and this is the whole issue about quality and safety, how do you create actionable data? Thank you very much.

CHRISTINE JORM: We've been working the last few months on the upcoming reports and, after considerable delays getting the data, which probably is a topic of some relevance to what we're going to talk about, we're advancing on now. What we've found is variation, but this is a different kind of variation. This is not in rates of procedures; this is in hospital-associated complications. This graph here shows cardiology patients, so that's patients admitted under a cardiology service with a whole range of cardiac conditions - heart failure, angina and so on and so forth - and the difference in the number of hospital-acquired diagnoses they then developed during their admission. So we're looking at cellulitis from their cannulas, we're looking at urinary tract infections, we're looking at pneumonia. The difference was startling. There was a 25% higher chance of a complication at Australia's worst hospitals compared with their best and we're expecting, when we analyse more of the data, that it's going to look like this all the time. Again, this is a model where hospitals can be accountable for what they do and routine data is important. We're certainly not doing enough to reduce complications. This is a graph from 2012 to 2015 of hospital-acquired diagnoses and they're cumulative, so we're looking at about 20% of patients get these problems and patients get usually about 1.5 of them. It's not going away, it's not changing, it's not budging and it hasn't budged for a long time.

So why aren't we doing better? Well, there's a lot of myth and tradition in safety and quality that has become codified and rigid and is just getting in the way of thinking logically about change. The question of preventability was a massive one, an incredibly distracting concept, could this thing have been prevented, and every study looking at preventability where they put experts in a room, the experts can't agree. To get accuracy you have to get a heap of experts, like 15 of them; pointless, expensive exercises. The second thing that happens is that the criteria for preventability changes all the time. We can prevent things that we once just thought happened, never better demonstrated than with Peter Pronovost's central line work where a rate of infection that was considered to be something that just happened was demonstrated to be possible to reduce to zero. So preventability is a moving space. The other problem with preventability is it introduces the potential for blame and once you have that people get defensive. "Oh no, it wasn't preventable." "My patients are sicker/my patients are fatter/my patients are older/my patients are different." So it's a problem. For those sick/fat/old patients, it's actually hard. Someone comes into hospital with renal failure and a whole series of other comorbidities and maybe their fluid balance wasn't managed so well by a junior doctor who wasn't so well supervised. Well maybe, but they were pretty sick. The effort to disentangle the causation is and has been unproductive, so the focus needs to change to the overall risk of patients getting complications and move away from what was really the blind alley in error theory that health went into.

The second problem we've had is the normalisation of harm. The lack of comparative data means that people just accept a rate of harm. They have no sense of what their own rate is, for a start. People can't remember something that happened in less than one in a hundred patients and it's probably even less

than that. If you don't remember 50 patients ago you probably think you don't have the complication, but for patients we're looking at trying to achieve much lower rates of complications than that. Then comparative data is just simply not available, so hospital boards and management don't know how they're going compared with others, clinicians don't know how they're going, and they don't know if they're improving, they don't know if anything's changing or anything has changed over time. The modern thinking in safety is really being driven by Erik Hollnagel, along with some of the other thinkers like Sidney Dekker, and it's this concept that he's come up with and he's titled it *Safety II* but it's called other things. It's very popular in industry now and organisations like the mining industry and Qantas are starting to turn their safety programs. The idea is that rather than focusing on individual harm, the emphasis is on reducing risk by improving system performance overall. So what you study then is different. You study normal performance and frequent events and you monitor all outcomes. We know when a patient is harmed it's because everybody was too busy or the communication systems didn't work well; there's a whole constellation of chaos in the healthcare system that staff normally manage to guarantee safety through, but it sometimes gets too much. But studying what happened on that exact day is not nearly as useful as studying what happens every day, what are the impediments to work every time, so *Safety II* thinking has implications for what we do with data and how we think about data.

Here's some of our recent analysis talking about the frequency of problems in the hospital. What we have here is the classification of hospital-associated diagnosis (CHADx) and for multiday medical admissions 20% of patients acquire one of these things they didn't come in with. The Commission on Safety & Quality has a much narrower list of hospital-associated complications where they focused on preventability and crunched the list down and only 3.5% of patients get those. The nationally reported list, the sentinel events, is a sad little collection of things which 0.1% of patients get. But when you look at this list and when you consider what I just said about *Safety II* it takes you to realising that the sources we need are big sources, because if we're studying big harms, common events, we're going to be much more effective improving systems and reducing the rare extraordinary events, so it changes the way you think about your data sources. We have all kinds of safety and quality data. Once upon a time it was all about incident reports. We've got patient-reported experience measures, we've got registry data, we've got death audit data, patient outcome data and accreditation, and I'm going to talk a little bit about all of these. But if the data doesn't help us improve, we're just goggling and there still is a lot of goggling with all these data sources. Patients are particularly poorly served. We don't have as much shared decision making as we should, in part, because clinicians are actually not very good at it because they don't know the risks, they're not in possession of the data they need to engage in it, and there are just a few comparative materials around for patients to make choices about where they go or who they see.

So it's not a good state at the moment, but I believe we can do better. There are potentials for all these data sources to become better - and by that I mean more actionable - and this is about a conversation that moves away from one is better than the other. They all have strengths - you'd be happy for any superhero to drop in when you're in trouble and I think a lot of the debate between the fans of one and the other has been extremely unhelpful as well. So what's "actionability" about? It's about four things. It's about soundness: it's whether people can trust data. It's about relevance and relevance is the one that is often really ignored: can you do anything with the data, can it help someone make decisions and is it relevant to anybody? The next one is accessibility: is the data in the hands of the people that can use it? Then, finally, understanding: can the data's insights be readily accessed by the relevant stakeholders? We've got a lot about all of this in the upcoming report, so this is just a little cook's tour of some of the ideas. "Actionability" is about helping patients. There wasn't an alarmingly large number of photos on the internet of superheroes terrorising children in children's hospitals, but I actually thought

it was nice to keep remembering that because data can seem like a dry old topic, but it is about improving the outcome for that child in that bed and every one of those data sources should be sharpened and toned and enhanced so it can do it better. Now, every one of them has its strengths and weaknesses. Failure to attend to measurement weaknesses makes data flawed, but it also makes our health system desperately inefficient. It's wasteful. These things are expensive, data collection is extremely expensive, and for us to be just doing the same old thing that never does any good for anybody is just throwing away money. It is an opportunity to do better.

Incident investigation reports, we do a lot of them. In New South Wales alone 140,000 incident reports a year, over 600 RCAs (root cause analyses). It's a local management tool. There is no data publically released in any state in Australia on any of this work. The quality of some of this work might be a bit doubtful. We certainly know counting incident reports isn't helpful, but one would hope for all the man hours that go into RCAs that they would actually be saying something useful. Perhaps they're not. There's some American research that's suggesting that maybe we need to be rethinking those processes, but at the moment this is wasteful data. This is not actionable, it's wasteful, and we can do better. Death audit data, we've got surgical, anaesthetic and maternal. The surgical death audit is a really interesting process in Australia. It's fully comprehensive, all deaths nominated by surgeons are assessed by a peer, then they get a second line assessment, and it appears that it's had a significant role in reducing surgical mortality. It's a correlation, it's not a causation, but it seems to have been very impressive. But there are problems. The surgeons will write a report which they'll share with the other surgeons and publish on the RACS (Royal Australian College of Surgeons) website. They're not secret saying things like "there were problems because there were only junior doctors in the hospital" or "the radiology wasn't onsite and we couldn't get the CT scan done". It sits there on the RACS website. It's not reaching administrators in a kind of way that might drive action. It's not reaching other members of the hospital staff.

In the case of anaesthetists the problem's worse. They have this narrow cone-down collection where they only look at deaths directly relate to anaesthesia or within 24 hours, so they look at a tiny set of deaths. We've been looking at bariatric surgery as one of our cases. There were seven patients in our dataset who died after bariatric surgery and only one of them died within 24 hours, so the other six patients having elective operations would not have had any anaesthetic review. Anaesthetists are experts in perioperative assessment, perioperative optimisation, and it's madness. Every elective patient who dies in hospital should be having an anaesthetic review. So this is an example of ways that our existing data could be ramped up, revved up, go and tell us a lot more things. Accreditation data, we debated about whether we included accreditation in the data section, but there's a heck of a lot of it. Again, a fantastically enormous amount of material is collected on every hospital, but it's not publically available. You can get the full public accreditation reports for medical schools and for nursing homes. I have no idea why they're not available for public hospitals. They're not going to ever be very user-friendly, but sometimes it's a matter of a secondary analysis by an academic or a journalist that'll point to the things that need attention. That's okay, not everything has to be designed for the consumer, but to have them in a secret world is just bizarre. You've got to also raise doubts about the soundness of the process. The substantial variation we have in patient harm, substantial variation in the appropriateness of procedures and the recurrence of scandals, like most recently Djerriwarrh in Melbourne, in fully accredited hospitals leads to doubt about the power of this in any fashion for assuring safety and quality or measuring it.

Patient-reported outcome measures (PROMs). I had to put Wonder Woman there because they're the big emerging thing. They're expensive. They're really exciting because they've got that potential for shifting the priorities in care right around to what's important to a patient, so patients can learn from the experience of other patients. The international research on them is currently really weak. It's all about enthusiasm and "we should". One of the big problems is the issue of setting up responsibility for taking action on the results of PROMs. It is completely unclear who would do that, so they're starting to fall in the goggle category very easily, so the patient doesn't walk so well after their knee replacement and is sorry they had the operation, who's going to do something about that? It's actually a problem and a much bigger problem when you're looking at disease based PROMs, looking at the outcome of a patient with cancer or rheumatic disease who's got social problems, who may be suffering from depression. You can collect a lot of stuff about how they feel, but when we don't have a mechanism for someone to take responsibility for action there's a problem with setting collections up. Patient-reported experience measures (PREMS). They're ubiquitous, we collect them in all hospitals, but they're not always accessible - in Victoria, to get the state results you need a Department email address, so the public cannot access the results which is pretty weird - and the soundness is very limited. Only in New South Wales are they standardised and patient-reported health status, language and age account for 20% to 30% of the variation in experience. So experience measures that are not standardised are useless because you don't actually usually have 20% to 30% variation between hospitals. So that's kind of horrifying when you think of all the thousands of these surveys and the millions of dollars that are going into these around Australia, but they're not actionable, they're not granular enough. Often the results are published at hospital level. Patients are treated by teams. Teams are interested in the experience of their patients in the respiratory ward or wherever else it is and this high level data is very hard for staff to use, very hard for patients to use to make decisions.

Clinical quality registries are powerful kinds of things. When they're good they're very good, but they're not all good. They can collect detailed information and return it to clinicians. You can have process measures, outcome measures, and they can follow patients longer term. There are some really good registries in Australia. The soundness though is impaired by their voluntary nature and there is a distinct lack of coverage of many registries, even though they purport to speak for the whole cohort of patients. Not that there is any need necessarily to do that either. Sampling has been a normal part of scientific methodology for a long time, so there's a sort of strange desperate effort that we've got to get them all as if that was going to make it better. They don't get them all anyway and the gaps are often in the area of private hospitals or particular states in Australia, so there are problems with their soundness. Their relevance can be limited by the special interests of the people running them. Surgeons are interested in surgery and not necessarily interested in the other things that happen to patients, so you're predetermining your complications in registries to a high degree, which is going to be less and less relevant as more and more of our patients have multiple health problems. They're not going to help us make decisions for those patients. Finally, feedback is not timely and accessibility is typically limited to the medical stakeholders, in particular there's a failure to release data to government, who do most of the funding and there's a reasonable desire to know what's happening. It's sort of a difficult area. We've done a piece in the report looking at a whole lot of registries and rating them on their cohort coverage, the nature of their data, their public reporting and whether they have a proper feedback mechanisms to the providers, and it was a bit of a dog's breakfast with dots all over the place, very messy.

I'm going to use a specific example now and talk about the bariatric surgery registry. This is data contributed by bariatric surgeons, it's got sponsorship - that's its funders on the right, always interesting in itself, a bit like a cycling jersey - and patient consent is required. The coverage of their field is very

incomplete. When they looked at MBS data they get 40% to 70% of the relevant procedures. I don't know which 40% or 70%, you know? It's interesting. They look at types of numbers, they look at defined adverse events, they look at three adverse events, just three - oh, and death - unplanned return to theatre, unplanned admission to ICU and unplanned readmission, and they got a rate of 2.2%. Now, remember that 2.2% because we're going to look at bariatric surgery in a different way in a minute. That sounds pretty safe doesn't it, 2.2%? They measure patient weight loss at one, two and three years and change in the patient's diabetes management, and those kind of long term things are things registries can do. They don't seem particularly good at it, they're losing a lot of patients in follow-up, so these things could be good but often aren't. I'm going to talk about routine data briefly. Stephen presented work based on routine data, the early slides I did were based on routine data, and that is the material that is coded from every medical record from every hospital patient that was admitted in the country and discharged. There are concerns about the soundness which relate to the quality of the medical record: if it's not written down it can't be coded and there are coding errors as well, so that's one of the limitations of this modality. It's largely inaccessible to patients and clinicians and it contains so much information that lack of analytical expertise limits understanding. In order for people to use it more it needs to be released in a form where it's classified, where it's cleaned, and where it's actually made available without a vast sum of money and several years waiting.

So, what can routine data sources tell us about bariatric surgery patients? We just ran this analysis today. Remember the 2.2%, because I want to suggest to you that, flawed though is it, it actually can tell us a lot more than we realise and it's right there, we pay for it, it's part of our everyday data landscape. So this is the safety of hospital care for bariatric patients by institution and you can see the significant and quite marked difference in the risk of complications for those patients between institutions. This is the length of stay for these patients and when we looked at the ones that were in for longer than a week we found that there was an excess of patients having gastric bypass, which is a bigger procedure than some of the others so they're the ones that end up stuck in hospital. So this isn't about criticism of surgeons or surgery, this is what can we learn from the data or resources we have about what happens? Here are the complications across all bariatric surgery patients. About 10% got a complication and the big one was gastrointestinal complications, not surprisingly, it seems perfectly reasonable, but when we looked at the patients who were long stayers their complications became quite different. They had much fewer gastrointestinal complications, many more of them were infections; many more of them had falls. So you're able to dive down with this data and find out different things. Then if we look again at specific infections, so this is our long stay patients, this is what happened to them: wound infections, pneumonia, mycosis, gut infections, sepsis, UTI, respiratory infections and implant infections. So suddenly we see the richness of this resource for diving down and actually understanding what happens to a group of patients. We just picked a group, bariatric surgery long stay patients, and found out how they were different to the ones that didn't stay long. That gives you a lot of clues for actually trying to improve the safety for those patients.

For all the data sources that we looked at there were three standout things that needed to happen. One is that we needed to improve data linkage processes and that's a whole topic on its own, but something we're just not doing enough. The second one was to extend reporting to all relevant stakeholders. All of these data sources were locked away and just shared with some people and there was much opportunity to do that. Finally, to report outcome data at hospital level because that is a level at which people can make decisions, GPs can make decisions, patients can make decisions, and it needs to include private hospitals in a way that they're comparable. PREMS, PROMS, any other measure should

be the same for public and private so people can make real choices. So they're some of the directions and ideas that we're hoping to include in the new report. Thank you.

STEPHEN DUCKETT: Thanks very much Christine. We've now got half an hour for questions.

AUDIENCE: I'm really interested in this idea of actionable data, but the decay in actionable data. When you were speaking I was reminded of Goodhart's Law which states that as soon as a government or organisation attempts to regulate or monitor any set of indicators they immediately become unreliable indicators of performance. I've seen that in my own experience in various institutions and I'm wondering how you stop indicators from becoming meaningless once people start to use them as indicators of performance?

STEPHEN DUCKETT: So I'll start. I want you to reflect back on Christine's presentation and that last bit about bariatric surgery and this was putting that information in the hands of bariatric surgeons so they can learn more about what was going wrong, they can drill down into what were the key problems, and they can identify that wound infection was one of the greatest of that class of things that were going wrong. Then you say well, the people who are benefiting from the data in terms of their ability to understand it and to use it are the very same people who are recording these things in the record. My experience over a long period of time has been the more data are used locally, the more useful they are centrally because the more interest the local people who actually put the data in have to get it right. That's the first point. The second point is there are errors in this data already, but generally the errors that we're seeing in the data are conservative errors and that is things are not being recorded. So if you look at comparisons of, say, the routine data and special analysis type data they always say, "The routine data doesn't include stuff, it doesn't include this". So what we're seeing here is an underestimate of everything and if you think about it, it still shows terrible things happening so if that's an underestimate, does it matter if that's the error that's happening?

So I think you're right that there is the risk that the more you set targets and so on they will be gamed, people will try and avoid them or whatever, but the story we're talking about here is making the information available to all stakeholders, making the information available so it can be used locally.

CHRISTINE JORM: I don't think it's quite answering your point. Can you give us an example?

AUDIENCE: So one simple example of, let's say, responsiveness to patients, that sort of thing. So let's say there was a ten day window in which people had to be responded to after a procedure or something like that people would just be sent letters with nothing in them so they can tick the box to say, "Yeah, that happened".

CHRISTINE JORM: Okay. Well, I guess there certainly will always be gaming when people set unreasonable standards that people can't manage to achieve in their work life, whether it's the Veterans' Administration or Canberra Hospital; the scandals go on around the world. I guess the focus here is actually providing better tools for people. You're never going to have standards around the surgical death audit, but you can get the results of the audit more clearly positioned so it can influence administrators and the people that are responsible for staffing hospitals, that kind of thing. Enhancing and things we're doing already so they're able to communicate better. Thinking harder about what decisions people might get to make with the data and whether they can make them. We've written an appendix in the report on numeracy because people's numeracy is often not very high and that includes

clinicians; there's this assumption that they'll get the numbers and they don't. There's quite a lot of research on the range of abilities there, but giving people better tools so that patients in particular, who have a right to know what's likely to happen to them, can actually investigate and find out and then make appropriate decisions. Stephen's more regulatory than me, but we're focusing primarily on improvement and that idea of the high performers helping the low performers and asking the "so what's happening here?" type questions with better data.

AUDIENCE: A lot of the things you recommend doing with the data seem pretty obvious - make it accessible, reduce how inaccurate it is, get it to people who need to know about it - and I wonder if why that's not happening is connected to the fact there are so many different types of data or the hospitals are overburdened. Why are these relatively obvious-looking things not being done?

CHRISTINE JORM: People are very precious about their data and they're anxious about sharing it. Remember, when things go wrong in health it brings state governments down. There is a concern about what it might show and the scandal that might emerge, so you end up with collections that really don't do a lot of good.

STEPHEN DUCKETT: Also, there's always a balance. Does a government want to lift the lid on what's happening in the hospital system? So 25 years ago we didn't publish sensible information about hospital waiting times, for example, and over the last quarter century we've seen much more information published about performance of state governments and hospital performance in particular. This is probably the next frontier of that public administration change and it's partly that we didn't have the tools to do it or there's been a development over time of the ability to do all this and there's probably more expectation that it be done.

AUDIENCE: You started talking about waitlists but I'm wondering if you want to make a comment about the way our public Emergency Departments (EDs) are measured in their performance based on just times - time to be seen, time to discharge patients, ambulance off-load times - and whether we are measuring that data because time stamps are the easier thing to measure and how that relates to performance?

STEPHEN DUCKETT: I can remember when we first started to measure and publish ED wait times. I was in my office here in Melbourne and one of the senior emergency physicians in Melbourne, a guy called Joe Epstein, came to see me and said, "All of the focus in public hospital policy is on elective surgery waiting times. It's just as bad in the Emergency Department and we're getting squeezed out!" So we talked about how could we balance the concern about performance in elective surgery with concern about performance in waiting times and we ended up with that measure, which may or may not have been a good one. However, when I think about these things I always think about the health system trying to balance three objectives simultaneously: access, measured by elective surgery or ED waiting times, quality and cost or efficiency. No one of those will ever adequately capture performance, so you're right.

CHRISTINE JORM: Also, it's particularly unfair the massive scrutiny of ED throughput where in other parts of the hospital people loiter around pointlessly waiting for things. So it's not fair.

AUDIENCE: I've got two questions. The first is you both flagged accessing data from the private system as a problem in your respective presentations, so I was just wondering if you could speculate on how

best to get access to that data. Secondly, I think one thing of concern when we're talking about actionable data is the lead time from when we discover that there's a problem to being able to action that data, so I was wondering if you could comment on strategies that could possibly decrease the lead time from when you identify a problem to actioning it?

STEPHEN DUCKETT: The private hospital data is interesting. It's got a number of components to it. We actually have the private hospital data in our dataset, but it's as if all the private hospitals in Victoria were one hospital, whereas for public hospitals we've got individual hospitals. We don't know the names of any of them, but they're separately identifiable, so the data provision to us has been restricted. What we're saying though is the government, when it's publishing information about hospital performance, in particular about quality performance, should publish it about public and private hospitals in exactly the same way. So there shouldn't be one set of rules for public hospitals and one set of rules for private hospitals. There should be the same set of data, the same metrics published for both public and private. The government collects that data and can do that now, it's a matter of political will. It's shrouded in all sorts of other arguments like "commercial in confidence" but "commercial in confidence" means there's competition and with competition you want better information, so because it's a commercial entity that's an argument to publish, not an argument not to publish, but anyway.

With respect to the lags, this is I think partly a technology issue. If you listen to Euan Wallace, who's the Head of Safer Care Victoria, he'd say, "Look, just last month we published the surgical audit data for the most recent annual report and we published the three previous ones" because they just were hopeless in terms of the lags in publication. So it's partly a matter of will, partly a matter of technology, but you can start releasing data on a monthly or quarterly basis, which is what they should do. When I went to Queensland a number of years ago they were doing their analyses annually and with about a two year lag. We changed to doing it quarterly with a three month lag.

AUDIENCE: I'm really interested in the other end, which is when you've actually collected the information, then translating that into actionable practice. We recently conducted research with over a thousand nurses across Australia attending educational events and asked them several questions which were about intention to change practice as a result of the new knowledge they've gained. There was a remarkable, if not stunning, consistency across the country in the responses of these nurses. It was open-ended questions and it came down to about five roadblocks and they're really interesting, involving things like confidence. Having the confidence to actually put into practice new knowledge was a real block and it came over and over again. As I say, there were about five in all. So I think the idea that once you know is one aspect, but actually then translating that into meaningful action and following through is non-trivial. We could see these, it was almost like road bumps along the way. I don't know if you want to address that, but I think that's an elephant in the room.

CHRISTINE JORM: You're quite right and there's been some very good research from the UK that's found just that, particularly studies where they've got the nurses involved and the patients submitting PROMs and PREMs. Great response from the patients and then the nurses are like, "You want us to actually do something?" They're all for the project, but it's that next step of making change that didn't feel in the gift for all kinds of ways.

AUDIENCE: If I could just say, it's not a one-way street. What the nurses identified was that when they knew something new they then needed enabling. It's almost like a vortex needs to open up to support that new knowledge through and the notion of the team is a really unstable notion. These nurses are

being expected to work with different teams consistently, with different people all the while, and so they're having to reframe their knew knowledge continually and express it in a manner that gives meaning and actually is effective. So I think that the whole health system is becoming much more complex, certainly the data that we're seeing, which is very interesting because when you ask people what they think and then you look at behaviour there's a massive difference in what's happening there.

CHRISTINE JORM: It's a really good point you've raised. For any person, if they're presented with data that they do not have the ability to act on they just suffer an amount of emotional dissonance. It's actually effectively unpleasant for them, they'd rather not know if they can't act. So it's a really good point you've raised. Making the data actionable in all kinds of ways is one thing, but providing the kind of supports in terms of organisational structure in terms of sensible central safety and quality bodies that are not just producing a million pages of rules and guidelines, those things need to happen and we haven't been so good at that.

AUDIENCE: I've just got one other comment there. One of the blocks was policies and procedures in that they would learn new information and then they would go back into their environment and there were all these bolt-ons that were there that made it impossible to actually change.

STEPHEN DUCKETT: That had been rigid for the last decade, exactly.

AUDIENCE: In terms of adverse events, I'm wondering whether the environment in which they happen is taken into consideration. If not, I would think that it needs to be because the adequacy of the facilities is a contributing factor.

STEPHEN DUCKETT: So the question was about to what extent do we take into account the design of the facilities in looking at the adverse events rates. There is now both a substantial literature about effective hospital design and safe hospital design and, in fact, quite an interesting controversy in the literature about whether single rooms are the appropriate policy direction, the argument being most obviously that if you share a toilet you're likely to share some infections, for example. We don't do that because our hospitals are not identifiable. We're not allowed to identify the hospitals so we're not able to do that, but it's a good point.

CHRISTINE JORM: It's the kind of thing that might be able to happen in the future. I mean, one of the things we haven't said is that Australia is way behind other countries in terms of public reporting. The amount of information you could find out about UK hospitals is just astonishing and it's all identified. No drama, there's just screeds and screeds of really interesting stuff about how they're doing compared with their peers and so on and so forth. US public reporting is far more advanced and it's not often the US is more advanced in healthcare. Australian public reporting is really lagging behind.

AUDIENCE: Most of the presentation focused around hospital care and what's happening there, but in the next 10 or 20 years I'm seeing a lot of home care moving in to a much greater extent. What happens then is that the offerings are going to collect more data with respect to what is happening at home. I just want to know how that would relate to what's happening in hospitals and how that improvement happens afterwards moving into home care?

STEPHEN DUCKETT: We've focused on hospitals partly because they're bigger in terms of the spending in the health system and partly because the data are there. The data available about primary

care is very, very weak and the measures of quality and safety in primary care are really, really quite complex. That's not to say it's not an important issue to deal with, but we're not equipped at the moment to do so. One of the things we're doing later this week is going to Canberra to argue there ought to be more data about what's happening in primary care to actually try and understand what happens in these areas, partly for the reasons you're talking about.

AUDIENCE: Is part of the problem that the damage that has occurred from bad outcomes, let's call it that, from hospital experience is not recorded? So you have a patient who has a bad outcome and they've become dependent upon their family who can't go to work etc. and we're not quite correctly recording the full cost to society of a bad outcome. Without that, you're not going to get the driving force for reform, because what you're saying is we need to record a lot more detail and we can see the economic and societal damage caused if we record a lot more data and therefore actionable data.

CHRISTINE JORM: It's a really good point. The story of someone with a complication, multiple readmissions, long term antibiotics, loses their job, becomes depressed, loses their spouse, I mean, these are normal adverse event stories and you're right, the economic cost of them is tremendous. I guess where we're going today is just some practical things that we can actually do better with what we've got already, because we're not using what we've got already. Your point is really well made and when we have a better joined up data system, if we ever do, in Australia between health and social services, which we ought to be able to have, that will help a lot, but it's never going to go into issues like the cost of a depression that follows on from an adverse event. So it's a really well made point.

STEPHEN DUCKETT: The first point in Christine's presentation was data linkage and that will help with that, but it won't address it all.

AUDIENCE: I'm interested in whether hospital level data is fine enough and whether individual practitioner level data would be a more interesting thing to look at, especially in a world where it's my perception, at least, that clinicians don't see themselves as falling under the clinical governance of hospitals, they see themselves as individual practitioners?

STEPHEN DUCKETT: A really interesting and complex question, so I'll give you two alternative answers and you can choose which answer you wish. The first answer is no, it is not appropriate to do individual clinician reporting because essentially what influences these sorts of things, including infection, is not only the individual surgeon, for example, but also the nursing team and all the other people, the cleaners and everybody else, who surround that patient in that hospital. Secondly, rarely does an individual surgeon do enough of something to be able to say with confidence that they are hopelessly better or hopelessly worse or whatever, so you can have that answer if you wish. Another answer you might like to say is this, that in the private sector an individual surgeon is taking individual responsibility and giving you an individual bill for the service that person has rendered. If they are giving an individual fee then they are taking individual responsibility, so we ought to publish individual results. So you choose which answer you want.

CHRISTINE JORM: I'll give you a third option. In the US now there have been a couple of really big data linkage studies that have found significant provider level differences. There's a big one looking at obstetricians and gynaecologists where they looked at their practice over about five years and then related it back to where they'd done their residency program. There have been others where they've looked at medical school and associated that with failure to or likelihood of practising evidence-based

care. So even where you're looking at maybe you don't do a hundred of a procedure, there are starting to be some techniques where you've got enough data about all the patients you see where I think we'll eventually be able to have individualised measures. Because that's what matters to people, something they feel that they can alter, that they can see. I think a lot of clinicians want that and I think there's a future for it too.

STEPHEN DUCKETT: One of the interesting things about this is there is a balance of interest. There's that the individual doctor has rights to privacy and if there are complaints made about them should you tell the hospital board and so on, versus the rights to future patients and the rights to individual patients. My view is that we've erred too much in our policy settings in terms of protecting the interests of the individual doctor and not enough in terms of protecting the interests of the future patients, but that's a judgement call about how policy settings are made.

AUDIENCE: Most people in the audience would probably know that in the US for a lot of adverse events that are deemed preventable they've stopped making payments to try and initiate quality improvements. Is there any outlook that the data will become sophisticated enough so that an organisation that, say, is an outlier for arthroscopies for arthritis you might be able to say, "Well, you've done X amount for arthritis" and withdraw their MBS payments to try and act as a deterrent?

STEPHEN DUCKETT: Yes, that is possible now. You could do that now and in terms of the US experience, Australia is moving in that direction over the next couple of years as well.

AUDIENCE: In terms of the clinical change that this could potentially drive, where do you see peak bodies as having a role in supporting that change or affecting that change?

STEPHEN DUCKETT: If you think about it, a lot of the stuff I presented was the evidence about what something works and what not. If you think about the bariatric example that Christine presented, the professional body had identified what were the three indicators. One of the critical roles of professional bodies should be what are the measures that we should be using to compare the performance of our members, for example. Many of these bodies claim they're interested in clinical standards and they should be saying," This is how we're going to measure that" for example.

CHRISTINE JORM: Yes, I think it's time for them to make a choice and it's a really important time for them to do so. You look at some of the work that's gone into the Choosing Wisely movement. Some of them stepped up and did decent things. I know some of the colleges produced this pathetic list of things that they never did anyway, "Oh, we're not going to do these ones anymore". It was such a failure of engagement with it and I think it's interesting times for the colleges as to whether they choose to take control or become a much less relevant organisation, as they have in the UK. They lost power there, they lost control, and I think it's that point for Australian professional bodies.

STEPHEN DUCKETT: Good point to end on, a nice challenge there. Thank you very much everybody for coming out on a cold night and thank you for the questions as well.

END OF RECORDING