Assessments in the new curriculum
Mr Rehan Khan

So, I mean, assessments are a large part of what we do and we have had to radically rethink this because of radical redesign at the end of the GMC. So, we're not unique amongst Royal Colleges in having to think quite hard about assessment, because actually, they're all trying to redesign things at the moment, and there are some new principles which the GMC released through various documents, along the lines of having a congruent programme of how you do assessments.

Within this are some things which, actually, it turns out we were always slightly ahead of the game, but we didn't know we were. Although we might not, day-to-day, use the phrase 'critical progression points', everybody in this room knows what the stakes are for, let's say, an ST2 ARCP. So, we know that because we're all involved in education, but actually, we know it because the next day they're a reg, and we're thinking, "How am I going to support them in that, in some ways, different role with different responsibilities?"

There's a lot more emphasis from the GMC about the idea that assessments that you do have to be assessing what they're supposed to assess and have to be consistent. We talk a lot about things being competency-based, but actually, thinking about things being outcomes-based is subtly different in some ways, partly because it's not just about whether you can do a particular technical skill. It's about the broader context, about how you do that skill together with the team, and partly because, at the moment, when you're doing those ST2 ARCPs, I'll have to use this phrase, tick box exercise.

So, I've done some educational supervisor forms in the last two weeks, because in London, a lot of ST2 ARCPs are just about to come up, end of this week or next week. Actually, it is a bit like you're looking at the evidence, but you're just checking it off, and then at the bottom, there's a wee box where you put an opinion. I'm never quite sure whether to simply write in that, "I recommend an outcome one," because I've shown all the evidence. There's nothing more I can say because of how it all works now, or whether there's more to it than that.

Actually, what I'm saying by that is you're ready to be a registrar in obs and gynae. I trust you with that. We were never quite sure if collecting the evidence in the way we were through a ton of OSATS, through a count up of mini CEXs and CBDs, which I may or may not have actually looked at, but I said I did, through a collection of TO2s, all of which have nothing under improvement needed or unsatisfactory. Yet I have a good knowledge of who this trainee is and what they're like, and what their likes are, what their dislikes are, and how I think they could develop.

I'm just not sure, and I wonder if you would share with me. You're never quite sure if the way we had it completely reflected what we knew about our trainees. I think we know more about our trainees than our current assessment portfolio suggests that we do. Therefore, there might be more, what we can do to help them.

So, trying to use the principles from the GMC, what I hope is we're moving forwards to doing something slightly better, and that, therefore, must involve the idea that we are allowed to make a judgement. We're not forbidden from doing that. It is perhaps not just a question of a reductionist count-up of whether or not a trainee has achieved certain very concrete things.

We are allowed, and we have always been allowed, it turns out, to make a judgement, but this has now become much more explicit, and we will talk about that, because, to paraphrase the late Stan Lee, with great power comes great responsibility.

Okay. So, the GMC talks about an integrated framework. That might include high-stakes exams, it might include workplace-based assessments, it might include multisource feedback, so that we can robustly evidence, ensure and communicate the levels of performance. At the progression points, which, inevitably, are annually, but in obs and gynae, we've got some which are slightly higher stakes than others.

So, at the moment, we are not bad at this, but it might repay being slightly analytical about how we do this at the moment. So, if you use the knowledge, skills and attitudes/behaviours framework, for knowledge, I think there's no doubt that we use high-stakes exams. We'll have a logbook, which, at the moment, is very tick box related, and we have some formative workplace-based assessments, which crossover between knowledge and skills. Skills, I guess we predominantly assess through OSATS.

Then attitudes and behaviours, this is where it becomes more tenuous and difficult at the moment, because we have multisource feedback. We have very little else, and I think some of the things I'd like to talk to you about is how we're trying to open that out with the new assessment framework, so that there are tools to help us to think about this.

That very much aligns with what Alistair was expressing around concerns that, at the moment, for any doctor in postgraduate medical education, it's perhaps too easy to fail to support them if they are okay with knowledge, okay with skills, but actually, their developmental needs with attitudes and behaviours, because the assessment toolkits don't quite let you do it. So, I hope that with things like NOTSS and with a greater emphasis on reflection, we might be able to move that forward.

Okay, so we know about critical progression points. We have this in our current ePortfolio, and that's absolutely explicit where the little breaks are, between ST2 and ST3, between ST5 and ST6. So, that has to be reinforced. There is, as I said, a need for more emphasis on validity and the GMC to help us produce this diagram. I guess the important thing in the diagram is probably the thing in the middle, which is what's the purpose of the assessment. Then you can, from that, have an overarching argument and analysis based on that assessment, to make the case that the programme of assessment has validity, okay?

With an outcomes-based curriculum, you have the opportunity to think about it in terms of things which are generic to all doctors, and then things which are really specific to obstetricians and gynaecologists, or specialists or sub-specialists in obs and gynae. So, the way in which this has been presented to the GMC very much reflects that.

Okay. So, it is important to acknowledge, I hope, that change is sometimes scary. So, the groups of people, the individuals and groups who, in honesty, did all the work – because I did very little, I just helped to bring them together – they expressed the idea that change was scary, but they expressed the idea that that didn't mean you couldn't change things. We've not changed anything, I hope, for the sake of it. We were really, really careful to have evidence and thought and engagement behind what changes we've done.

So, an example of that is what's on this slide. So, we assess knowledge at the moment through exams and we assess skills using OSATs, and we have not fundamentally changed these ways of doing things. Although the nature of the curriculum means there are some subtle things which have had to change.

With the old training system, we had a very, very long, 19 module plus ultrasound curriculum logbook, of which some of it was curriculum, which none of us read, and most of it was logbook which all of us use almost every day. Because it has a lot of tick boxes in it, it gives you a knowledge framework. Now, part of what we realised is, by going away from that way of doing things, and having far fewer tick boxes, we actually had to have a different way of making it explicit, what the knowledge requirement was for the exam. So, that is, in fact, called now a knowledge requirement, okay?

It's been grouped into knowledge areas, and, in effect, that's a learning guide for the syllabus. We're very influenced from the RCGP who've done it like that for quite a long time. So, we had to make that absolutely clear, because you have to make it clear to people what they're expected to learn for the exam, but the exam itself remains the exam. The summative OSATS have been looked at, and are essentially fit for purpose and remain, although there was a massive debate, in some ways, resolved, in some ways, unresolved, about how many you need, but in the end, it has remained the same, and under a lot of guidance actually from our trainee committee.

So, these are our current and future formative assessments, and they're familiar. So, familiarity at the start is, I hope, helpful and important. So, it is still true that we use CBDs. It is still true, they help us to assess the performance of a trainee, in managing a patient, and give an indication of competence in areas likes clinical reasoning, decision-making, application of medical knowledge, in relation to patient care. It's still a way to document a conversation about, say, if a trainee's doing a case presentation, okay? It is still the case that we use mini CEX to look at the direct encounter with a patient history and examination, direct clinical reasoning.

We still and always will, I think, give immediate feedback at the time, and it's still a way to record it, but what we did think is that the way we do it now, it's got real benefits in principle, but it has impediments to making it of high quality. One of the impediments is, frankly, the idea that you get sent a blank form for something which might have happened two weeks ago. That is a big challenge.

We've had a lot of feedback on that, and we looked at a lot of other Royal Colleges and how they did it, and realised, with a little work through ePortfolio and a little bit of thought, that there was a different way to do it, which changed the way you looked at it, in terms of the learner being an adult. It's the idea. The learner actually owns the learning event, and, therefore, the person best placed to populate what the learning is going on on this, is the learner themselves and not the trainer, okay?

So, by putting that and then helping the trainer by making it clear that their role is to help the learner work out what the reflective element is, we hope that we can improve the quality.

So, we still have mini CEXs and CBDs. They are still formative, but by reformatting them, as we're trying through this pilot, I hope that the quality will be different.

So, this is a blank form, but the one important thing to note about this is, sure, the boxes are different, but when the trainee submits the form, the trainee themselves on their own has to describe the event and analyse it. It's got to say what happened, and actually, they've then got to say, "Why is this a learning event? What was the point of this? Why are they actually submitting this? What do they think they've learned from it? What learning gap has it exposed?" and then they can put their learning plan on it, and that's the document which goes to the trainer.

So, what we hope is this is different from a trainee sending something and then the trainer can't remember what it is, or can partially remember what it is, but doesn't know what it was that the trainee wanted to get out of it. The trainer can then read it and make their own analysis with recommendations for reflection, and then the trainee reflects on it and the item's finished.

So, what we hope is that that's something which the trainee can go back to in the future, and actually, it will have had value, quality to it, rather than running the risk of being an exercise that you do because you have to do it. So, that's one of the pilots at the moment.

All current, probably most, nearly all, current doctors entering our specialty went to foundation school. In foundation school, you've got to do multisource feedback – that's called a foundation tab, I don't know if many of you may have seen that – and in one keyway, it's different. So, you get feedback from lots of people, but before you can see it, the trainee's got to write feedback on themselves. That form is called a self-tab.

Now, actually, that is entirely, in educational theory, the way multisource feedback is meant to work, because the way in which you're meant to learn from it is you reflect having seen any discrepancies or similarities between what you thought about yourself and what the team thought about you.

That has been quite challenging in obs and gynae because we didn't do self-feedback in that way. Well, technically, we did because there's a box at the top of your induction appraisal form. This is the bit where your trainee writes how it's been going. This is the bit where, when you generate the form, at that point, the trainer usually goes off and has a cup of tea while they make the trainee fill out that box, and then they come back and it says, "I think I've been doing okay."

Actually, to be robust about this and try to get the maximum learning out of it, if you ask the same elements that we asked the team, to the trainee themselves, you can actually have something very much in parallel with what foundation trainees do. So, this also is being piloted and will be coming, I hope, in.

Could you raise your hand if you've had a look on the ePortfolio, at the NOTSS tool please? It's pretty much everybody. Well, it came from the Royal College of Surgeons in Edinburgh originally. I think it's fair to say that we use this a lot more than any Royal College of Surgeons do now. I've yet to meet a surgical trainee who's ever used it actually, although I'm sure some do, but this is very important.

If we think about what Alastair said about GMC referrals in the first few years of CCT, if we think about this college and other colleges' work in the Each Baby Counts project, we think back to that video. A very striking video about situational awareness on the labour ward. If we think about the training that a lot of us have been going through in human factors over the last couple of years, medicine is changing, obs and gynae is changing, the way we're thinking about safety is changing. Actually, the way we're thinking about what a doctor is for has changed.

So, it cannot just be about the technical skills and the knowledge. It must be about more than that, but until this, we never really had a formative tool to even have a congruent conversation about this.

So, by having a way, I hope, to think about things like situational awareness, decision-making, clinical and non-clinical, teamwork, leadership, communication, I think this opens out the idea that the modern obstetrician and gynaecologist has to be proficient in these skills.

So, we debated long and hard whether there was a way to make this a summative assessment, a pass/fail, and we've come to the conclusion that this is a learning tool, and this is formative. It is not in the spirit or of the educational theory behind it, or in the spirit of how we want to use the tool to try and make it summative. This is a formative tool, but we want it taken seriously.

It will be mandatory, we think, because, whether you're an obstetrician, a gynaecologist, or both, you may be in the operating room over the course of a session, you may be on the labour ward of the course of a day. You will need to demonstrate these skills, so that we trust each other and so that, at the end of ST7, I trust this trainee to be my consultant colleague. That idea of trust, I think, is quite important. Every night, when we we're at on-call, eventually, unless you're resident, we'll go home because we trust the person to do it right.

We really feel that there's an opportunity here to have an assessment portfolio, which, in a way, talks about that trust, and we have a chance to think about that annually through global assessment. Now, global assessment has the potential to be controversial because it needs maturity on the part of the trainer and the trainee to get it right.

I think, with the best intentions in the world, over many years, we built training systems and assessment systems to have a type of objectivity in them, so that we could try and eliminate bias, actually. I think, historically, in postgraduate medicine, some of that bias was sexism and racism, and other protected characteristics, and this needs maturity and care, because the GMC, in their own documentation, use words which we might perceive as challenging. They use words like 'it's important that the educational supervisor should be able to write down their feeling about this trainee at the end of the training year'.

So, we all worry about this stuff. We are all educators. We worry about that stuff all the time, but as well as being educators, we are human beings. I'm sure that it would not be true to say that people don't have feelings about their trainees. I'm not saying we fancy them. I mean, feelings about their performance, but we are now empowered to make a global assessment. It does have to be evidenced because, by the end of a training year, there will be a catalogue of evidence as set out using the curriculum and assessment portfolio.

So, the evidence will be there, but it emphasises the importance of encouraging educational supervisors to do true educational supervision, and that probably means, for many, doing something differently to how they've been doing it. It probably means, even if it's for 15 minutes, talking to their trainees explicitly about their education much more commonly than we're currently doing, and writing it down. There's going to be an effort and I think there are bits here which are going to be hard, but I am hoping that, although it will be hard, that it will be better.