**Examiner Training recording**

0:01
Hello, my name is Anahat Daboom.

0:03
I am the Head of Assessment and it's my pleasure to welcome you to the Examiner Training.

0:08
This short presentation is divided into two parts, approximately 20 minutes each, so Please ensure that you take break.

0:16
I would also like to take this opportunity on behalf of my teams to thank you for your time and expertise you provide as part of your role as our OSKI Examiner.

0:31
In this session, we will start with describing and comparing varying Oski formats.

0:38
We will then look at the standard setting process, which is essentially how we set the pass mark.

0:44
We will also look into the requirements of the Oski examiner.

0:50
First of all, some background.

0:51
So we offer several courses at Queen Mary and the first thing to note is that we have a London and a Malta campus.

0:59
So essentially this means that MBBS programme is delivered on both campuses at the same time.

1:05
The same curriculum and students are undergoing the same assessment.

1:10
They have the same assessment and assessments in both campuses.

1:14
They sit the same written papers and they sit the same oskies at essentially the same time.

1:21
So as you can imagine, this comes with quite a number of challenges for reasons related to timing, special arrangements and of course people, members of staff involved in the delivery.

1:37
We have several cohorts of students.

1:39
We have a five year MBBS undergraduate programme in both London and Malta.

1:44
We have a four year MBBS graduate entry programme in London.

1:47
We have a three-year MBBS direct clinical entrance programme in London, A2 year postgraduate physician Associate studies programme in London and each of these programmes has an OSKI in every year of the course.

2:06
The next thing to note is that we have upon us the UKMLA for medical graduates.

2:13
The UKMLA is the UK medical licencing assessment brought in by the General Medical Council.

2:21
It's a national exam designed to demonstrate that those who obtain registration with a licence to practise medicine in the UK can meet a common threshold around 3 themes.

2:32
Readiness to practise, safely able to deliver person centred care and capable of managing the uncertainty.

2:44
Pilots for the UKMLA began in 2122 and academic year 2425 will be the first delivery of the UKMLA, including the CPSA.

2:58
So that means that medical students now graduating in 2025 must have passed the UKMLA in order to be provisionally registered with the GMC.

3:10
The UKMLA content consists of two components and Applied Knowledge Test, otherwise known as AKT.

3:18
This is a mixture of single best answer questions.

3:22
There will be two SBA papers lasting two hours each with 100 questions each.

3:29
So altogether there will be 200 questions that the students have to sit for the AKT and this is developed by medical school councils with the oversight from the GMC.

3:41
This also includes the standard setting for the exam.

3:45
The Clinical and Professional Skills Assessment or CPSA is the series of OSKI stations and this is developed and delivered by each medical school with a very close oversight from the GMC.

3:58
Now as part part of the requirements and recommendations, all OSKI examiners will require training before participating in the future Oski's and this is what we are here today for.

4:12
As I said, as I mentioned, this is part of the recommendations and requirements posed by the GMC.

4:23
So, so some background again, let's start with the Miller's pyramid.

4:29
I am sure that many of you are already familiar with this model.

4:34
So this is just a revision.

4:36
We have the bottom of the pyramid knows, then we have the knows how, then we have shows how and we have does and essentially the knows and knows how is where we look at the cognition and knowledge in our students and the shows how and does is where we look at behaviours, skills, attitudes.

5:00
And obviously, as we climb up the pyramid, we are increasing professional authenticity.

5:11
So in the nose we can use single best answer questions.

5:15
So that's all very well, but it's often a waste of testing time just to test Rd learn facts.

5:23
So what we would like to do here is also test the nose how using SB as and these are more complex single best answer questions where students have to apply their knowledge.

5:37
So we can actually assess both Nos and Nos how with the single question.

5:43
So it's not then just recalling the pure fact we then have shows how, which is the performance assessment of the Oskys and which is what we will be asking you to examine in your role as the OSKIE examiner.

6:00
And then we have the does.

6:02
Now this is usually on placement in clinical environment where students use workplace based assessments via log books.

6:10
They've given feedback and students then use that information to improve their skills.

6:17
But although it's the highest level of pyramid, it's not always the place where we can ensure fairness, broad sampling of cases, structure or systemicity for students in their assessment.

6:33
It may be highly valuable to ensure that students have gained adequate broad enough skills and competency they require in order to graduate.

6:45
So let's move to or let's focus on Oskies.

6:49
It is the objective structure to clinical examination and it was developed by Hardin in 1979.

6:57
So it's been with us for quite some time now and over the years it has been there had there has been some modifications and adjustments, but it is currently the best tool that we have to assess students for their clinical competence.

7:13
So why do we use oskis?

7:15
We use oskis because we can give the stations as much validity as we can in the what really is the artificial environment.

7:26
We have careful specification of content and we can test performance of clinical skills.

7:32
We also have the reliability, so we have the observation of wide sample of activities and a structured marking schedule.

7:44
We also want to ultimately ensure the fairness between the students, so each student has to perform the same task and the interaction between examiner and student is structured.

8:00
When we are thinking about the exam, we have to also think about the characteristics of the exams like validity.

8:06
There are really 2 main areas that we need to consider when designing designing of skis and the first is validity.

8:14
So the question stands, does the assessment measure what it is intended to measure?

8:21
The results of the assessment are appropriate for a particular purpose as demonstrated by the coherent body of evidence that evidence.

8:30
So the evidence for validity of assessment will help us in justification for decisions which we make on the basis of the assessment results.

8:42
So one of the most important questions is, are the assessment decisions and outcomes of that, of those results defensible, credible and fair.

8:54
It is important in both summative and formative assessments.

8:59
And of course, there are multiple versions and facets or subtypes of validity.

9:07
2nd is the reliability.

9:13
That's another test characteristic that we're just going to stop over for a second and consider.

9:20
So here we are asking how reproducible are the scores across the raters, questions, cases or different occasions?

9:30
Are we using the same ruler and is it a correct ruler?

9:35
Does the test consistently differentiate between students of varying abilities of the good from the poor, the poor from the from the from the top students?

9:47
So the reliability can be controlled somewhat by making Ausky stations either harder or easier.

9:54
So if we wanted to differentiate the abilities from different students, then we would probably make the stations with some difficulty.

10:03
We wouldn't make them too easy because that would mean that everyone would pass and you wouldn't be able to differentiate everybody's ability.

10:14
So you would make sure that there is a sufficient reliability within the exam to ensure that it has appropriate characteristics.

10:24
So question stations shouldn't assess the same content.

10:27
That's another thing.

10:29
However, the majority of stations may assess infection control as a small part of the station by would but would not make the main focus of every station.

10:40
In order to also achieve a decent reliability, we also look at the competencies that are being assessed.

10:46
So what we aim to do is to have a broad sampling across different skills to obtain adequate reliability.

10:54
We want to obtain a really good saturated picture of the student in front of us.

11:02
So we would use a range of cases or situations.

11:07
As a simple example, we wouldn't have 15 stations where we assess a student on how they take blood pressure.

11:14
This would not be a broad sampling, so lots of different skills.

11:18
We want to then give us a decent amount of high highly domain specific stations with adequate reliability.

11:31
In order to enhance reliability, we can also reduce the error variance and this occurs in several different ways.

11:38
So we have simulated patient training.

11:40
The actors that we use in our stations are trained and we make sure that they are consistent between each of the different students that they see, but also the different circuits that we use and also in different sites.

11:55
We already spoke about the fact that we have London and Malta, We have also examiner training and that's what we are here for today.

12:04
We have clear and explicit candidate instructions.

12:07
We have a structured marking schedule.

12:09
We have consistent equipment in all stations and consistent conduct of the exam.

12:18
We use two methods at QME.

12:19
Well and depends on the year of assessment.

12:22
There are two types of structured schedule that we could use in the OSKI and these are domain and a checklist.

12:29
As a school we have moved from checklist to domain for a number of years now in our year three and four exams.

12:36
But as of 2023 we moved to a checklist marking system for all of our schemes.

12:48
Now I would suggest that this is a probably good time to take a break if you feel like.

12:54
If you got to this point on from the next slide, we will be starting the second part of the OSKIA examiner where we will be looking closer at the examiner's role.

13:11
Hello again.

13:13
Welcome to the objective start Structured Clinical Examination Examiner Training Part 2.

13:19
In this part of the presentation we will be looking closer at the role of the examiners.

13:30
So what we expect from you contributions made by you as an examiner means that the exam runs in a fair manner.

13:40
The role of the examiner in the OSKIES is to observe the performance of the student at the particular task, score accordingly to the marking schedule, and to contribute to the good conduct of the examination.

13:54
Examiners should not engage in conversation with candidates.

13:57
It's perfectly OK to say hi, but starting conversation with how are you, what are your summer plans or, or how did how did it go so far?

14:10
It's probably not appropriate.

14:13
Examiner should not make verbal comments such as well done, very good, fantastic, or Oh dear.

14:19
Examiner should stop assessing candidates when the timer is complete and examiners need to read the script to be familiar with the content of their station.

14:30
Example for when to ask questions or how to ask questions verbatim.

14:39
To continue our discussion we have to ask make a pit stop at standard setting.

14:46
So to look a little bit closer on how do we set the pass marks within a standard setting process, we have to use the performance based standard setting methods as this is the requirement of the GMC.

15:02
We can't use norm reference methods such as saying the pass mark is 50%.

15:08
In order to do this, we use what we call the borderline regression method.

15:13
This has been developed for more than over three decades and as oskies were being developed and has evolved into different types of marking schedules and different types of methods were used to generate the pass mark.

15:31
And we have now settled on the borderline regression method for large cohorts of students.

15:37
So obviously as the name indicates, it's a standard setting method using borderline regression method surrounds what is essentially the borderline candidate concept.

15:52
So during which during that process we have to ask who are they and what characteristic the borderline candidate might have.

16:06
So I'll give you a couple of moments just to think maybe about the borderline candidate, how this concept may look like, what behaviours, what characteristic might they have and you can choose the year group that you are planning to examine first.

16:30
Well, I'm interested if that task was easy or difficult and actually was there a person, an example of the real life student that you envisaged when thinking about the borderline candidate?

16:50
The borderline candidate is a theoretical concept.

16:53
It is very important to remember that examiners need to envisage the borderline candidate and it's a cognitive task.

17:03
In literature and evidence, it's described as a very difficult to perform.

17:08
It's a highly demanding task to create the conceptual image of the borderline candidate.

17:16
So the descriptor of the F1 on the first day may skew that image.

17:21
How?

17:22
Because as we are perceiving someone who already successfully qualified.

17:30
So that's the kind of philosophical underpinning.

17:37
They're not really the yes or no, good or bad kind of dichotomous concepts.

17:48
It's in the literature, it's been described more as a continuum, a spectrum or a zone where competent can sometimes overlap in a sense with incompetent and the anchor statements.

18:07
So descriptors in a specific concept are also based on that concept of the borderline candidate.

18:18
It is not your average student and we're thinking more about the exam performer.

18:25
So in my little literature search, I found a few descriptors that might help you in kind of painting that picture of your borderline candidate further.

18:41
So we said that we are not talking about the average student, we're talking about the borderline candidates or borderline exam performance.

18:50
What kind of characteristics?

18:51
We can describe them.

18:54
Literature uses several descriptors.

18:57
It's a hesitant candidate.

19:00
They might be insecure, they might be patchy or superficial, they might display limited reasoning or analysis, and they might learn by what.

19:15
So a borderline candidate might not have all of these characteristics, but they might have some of them.

19:22
If a borderline candidate has had all of these characteristics, then I'll probably say that they are on the verge of being a failing students.

19:31
But the borderline candidate who is patchy and their knowledge, for instance, may mean that they are good in other areas of, of, of some examinations or knowledge or, or skills.

19:46
So it's about mixing in these different characteristics to understand what the borderline candidate would look like.

19:52
They're not failing.

19:56
That might be the candidate you would like to see more of.

19:59
You would like to supervise them maybe a little bit closer, just to ensure that they are fine.

20:10
The borderline candidate for year five.

20:13
It's an important concept and again very well described in light of the CPSA and the MLAAKT exams.

20:24
So the borderline candidate for all Year 5 MBBS paper D stations is defined as the just safe foundation.

20:33
Year One doctor in week one on their first job after graduation.

20:39
This definition of the borderline candidate is consistent with a similar statement used as part of the AKT and the standard setting procedure used for that assessment.

20:52
The application of this term in practise starts with the instruction from the Senior Internal Examiner, the SIGH to the station writers.

21:01
So what would the performance of the Justsave, and I'm just highlighting you, the Justsave Foundation year one doctor in week one of your first job after graduation look like here.

21:16
When staff review existing stations or develop new ones, that's the concept that they are that is being repeated.

21:25
This definition is also included in the examiner training like today in the in multiple resources and you will be able to explore not only for your 5 but in general that borderline candidate concept later on in the workshop discussions.

21:50
So just again, to go back to the borderline candidate across all the years, this might be the person who is minimally competent.

22:00
And another word used in literature to describe that candidate is that candidate was inconsistent.

22:13
So we said that the standard setting is delivered through the borderline regression, which is statistical method.

22:21
So you're probably asking, well, what is my role in the standard setting as an examiner?

22:29
Well, your role in standard setting is to use your clinical expertise to judge the candidate's performance to first of all to allocate the global judgement based on the concert overall performance at the station and to remember the level of examination.

22:46
So is it U1U2FY1 as we discussed before.

22:51
So you have to remember that if you're assessing for final year medical student, then when you come into the examinations you will be assessing at that level each station.

23:06
So you allocate the global judgement based on the candidates overall performance at that station and that gives the candidate a global score.

23:18
So now we've got 2 elements.

23:19
In order to develop a regression line with the global score we have 4 criteria, good, pass, borderline or fail.

23:31
So for each candidate you will produce as explain and score and a global judgement and we ask you to provide feedback which is not used in standard setting process.

23:47
For the borderline regression method, you mark against the checklist.

23:52
There is an example of the checklist and you will give your global score and then we plot this data once the exam is finished onto the XY access.

24:05
So we put the good pass borderline and fail along the bottom axis and we put the score attained by the candidate on the side axis and we plot every single student.

24:20
We plot every single student from the graph.

24:23
We then enter a line of the best fit across all the data and where the borderline group meets the line of the best which this becomes the pass mark for the students.

24:42
So for all intents and purposes, you will not know if a student has passed or failed station when you are marking them because that determination is made after the exam, after we have all the scores completed.

25:01
So we want to know what the pass mark is until after the exams.

25:05
After all the students have completed the OSKI stations, we generate this graph to find out where the borderline meets the line of the best fit to give us the pass mark for that individual station that you've been examining.

25:20
So in a single OSKI, we can have anything up to 3000 different student examiner interactions, three 3000 points data points to construct the pass mark.

25:33
So that's a lot of data and that gives us a very robust analysis of what the pass mark for each station should be.

25:44
The discrimination facility of all sorts of parameters in the OSKI are also analysed at that time to ensure that the exam has been conducted at the highest quality.

26:01
Right.

26:02
So some summary, what would we like you to do is to contribute to the good conduct of the exam.

26:11
So that means please don't start the conversation with candidates.

26:15
You can say hi bye, but essentially keep your poker face and please stick to the script.

26:25
If there's anything that concerns you before, during or right after the exam, please report it immediately.

26:32
Report it to a Marshall circuit supervisor or internal examiner or myself if I'm around in, in your, in your circuit.

26:43
Any of these members of staff will be appropriate to take your report.

26:48
We will of course, act appropriately and as soon as possible.

26:54
This is a part of our quality assurance.

26:57
We will be reporting or any minor and major events during the exam at the subject examination board and we'll be informing our external examiners.

27:08
Now this is all anonymous, so please don't be alarmed.

27:12
We may ask you some extra questions just to understand what has occurred or what you've noticed.

27:20
We always appreciate that because that ensures that the quality and the fairness is preserved.

27:28
As much as possible.

27:32
Observe the candidate performance score according to the marking schedule please and don't forget the global score and don't forget the level of the candidate that you are examining at.

27:46
Submit yours marks as soon as possible and stick to that scenario.

27:51
If you have questions, you can always ask in between the students then or between the circuits is the best time.

28:00
But if you're unsure, there will be plenty of members of staff on site.

28:09
Now, I'm sure some of you who were not exposed to our oskies or to our mark sheets might be a little bit worried.

28:21
This will be covered in a workshop.

28:24
We will have an opportunity to discuss further some of the concepts like the borderline candidate or regression model and how you produce your skull scores.

28:40
You will have an opportunity to observe some cases and have a go at marking on the examination software you will have received.

28:51
You will or you, you have already received an ID and a web link to use during the session and then the pin will be provided during the session as well.

29:01
So I hope you're going to have fun during the workshop.

29:03
Thank you so much for your attention.

29:05
Thank you for attending and tuning into this presentation.

29:10
My e-mail is at the first slide, so if you would like to e-mail me with any comments or questions, please do so.

29:20
I will be delighted to receive.

29:23
Thank you and have a good day.