

Objective structured clinical examination – en etableret metode til at prøve lægelige kliniske færdigheder

Læge Anne Lindebo Holm & afdelingslæge Knut Aspegren

Syddansk Universitet, Det Sundhedsvidenskabelige Fakultet, og Sundhedsstyrelsen

Resumé

Lægeligt arbejde er en profession med videnskabelig basis. En læges kompetence afhænger af dennes evne til at integrere viden, færdigheder og holdninger. Det er derfor nødvendigt i lægeuddannelsen at anvende prøveformer, der tester interpersonelle og intellektuelle handlekompetencer. Objective structured clinical examination (OSCE) er en sådan internationalt anerkendt prøveform. Ved at opstille standardiserede og strukturerende scenarier baseret på kliniske situationer kan man på kort tid observere og bedømme et bredt udsnit af lægelige kompetencer.

Inden for dansk lægeuddannelse bliver termen »OSKE« brugt om forskellige slags prøver, som har det til fælles, at de foregår med flere eksaminatorer og med opgaver i rækkefølge. Vi mener derfor, at det kan være af interesse nærmere at beskrive den oprindelige prøveform, og hvad den kan anvendes til. Termen »OSCE« er engelsk og akronym for objective structured clinical examination. Man ser det ofte oversat til dansk som »OSKE«, objektiv struktureret klinisk eksamen. Men det engelske ord *examination* har i denne sammenhæng betydningen »prøve« og ikke »eksamen«. En OSCE er som oftest en del af en eksamen, på samme måde som f.eks. en »spotprøve« i anatomi er det. Da man som oftest ikke har vanskeligheder med at inkorporere engelske ord i det danske sprog, ser vi ingen grund til at bruge akronymet OSKE.

Prøve af lægelig kompetence

Lægeligt arbejde er en profession med videnskabelig basis. Det er derfor ikke tilstrækkeligt at beskrive lægelige kompetence som viden. En måde at beskrive den på er: vanemæssig brug af viden, færdigheder og holdninger i en professionel sammenhæng [1]. Når man skal måle lægelig kompetence, er det således ikke nok at måle viden alene, også andre elementer af den skal måles. Ved en revision af curriculum ved lægeskolen i Manchester [2] lavede man en analyse af de forskellige sider af en lægelig kompetence og fandt frem til følgende:

1. Viden (*knowledge*)

A. *Teknisk*: Om biomedicin

B. *Kontekstuel*: Om sygdomme på individ- og samfundsniveau

2. Færdigheder (*skills*)

A. *Intellektuelle*: For eksempel medicinsk beslutningsproces

B. *Interpersonelle*: For eksempel procedurer, samtale med patienter, operative færdigheder o.l.

En stor del af den lægelige kompetence består af intellektuelle og interpersonelle færdigheder. Disse »handlekompetencer« kan ikke testes på samme måde som viden. Derfor er det nødvendigt med andre prøveformer end kun test af viden.

OSCE blev oprindeligt udviklet til primært at teste procedurer med. I de 30 år, som prøven har eksisteret, er den blevet udvidet til også at kunne bruges til at teste mere komplekse interpersonelle færdigheder som samtalen med patienter og intellektuelle færdigheder som medicinsk problemløsning. At teste ren viden har aldrig været formålet med OSCE – til det er der bedre egnede metoder. Ligeledes skal man være opmærksom på, at man med en OSCE tester, hvad lægen/den studerende kan gøre i en testsituation [3], og ikke hvad lægen/den studerende rent faktisk gør i rigtig praksis. Der er udviklet forskellige metoder til at foretage en struktureret bedømmelse af den studerende/lægens virke i klinikken med, baseret på direkte observation af en tutor. Her skal nævnes to aktuelle: Clinical Work Sampling [4] og mini Clinical Evaluation Exercise (miniCEX) [5].

Anledning til at udvikle OSCE

De oprindelige anledninger til at udvikle OSCE var, at der var et behov for at kunne teste lægelige færdigheder og opnå en højere grad af inter- og intrabedømmerreliabilitet, end man kunne med direkte ustruktureret observation. Direkte ustruktureret observation af to uafhængige bedømmere giver som oftest forskelligt resultat grundet observatørens opfattelse af niveau, grad af træning i sin rolle o.l. Den samme bedømmer kan også vurdere to ligeværdige præsentationer forskelligt pga. forskelle i f.eks. køn, alder, tid på dagen eller andre konfoundere.

Historie og beskrivelse af metoden

OSCE-metoden er i stor udstrækning opfundet og udviklet af *Ronald Harden* [6, 7], som er endokrinolog og professor i medicinsk pædagogik og i mange år leder af Center of Medical Education i Dundee. Hans udgangspunkt var den engelske eksamen til de forskellige Royal Colleges, hvor den enkelte på

VIDENSKAB OG PRAKSIS | OVERSIGTSARTIKEL

en og samme dag gik rundt til flere eksaminatorer, *stations*, og således blev overhørt af flere personer. Siden er OSCE blevet udviklet i mange retninger og dækker nu ofte over en ikke altid gennemskuelig prøveform, der dog som oftest har flere ligheder med end forskelle fra det, *Harden* oprindelig gjorde.

En OSCE er således en række »stationer«. Hver station varer som oftest 5-10 min, hvilket kræver, at opgavens omfang tilpasses dette tidsrum. På hver station skal den studerende udføre en færdighed, som bliver bedømt af en og samme eksaminator vha. et i forvejen opgjort skema. På denne måde opnår man, at flere færdigheder kan testes inden for et rimeligt tidsrum, samt at alle studerende får samme opgaver og bliver bedømt efter samme kriterier. Hvor det er muligt, analyseres den enkelte færdighed i sine dele ned til helst et dikotomt udfald, »gør/gør ikke« (mest præcise bedømmelse).

Eksempel på en sådan testprotokol findes i **Figur 1**. Relativt komplekse færdigheder kan også bedømmes på denne måde, men her er der nogen evidens for, at globale scorer er mere pålidelige [8]. Prøvens indholdsvaliditet sikres ved at kombinere 18-20 stationer. Ved indholdsvaliditet forstås, at eksamen i sit indhold og sin sammensætning er afstemt i forhold til det, man vil teste med prøven, f.eks. kliniske færdigheder opnået under et klinikophold. I en OSCE, hvor man kombinerer flere stande/opgaver, kan man opnå både at dække indholdet og at vægte det. For at sikre at dette sker, laves der en indholdsmatrix for en OSCE, hvilket er en opstilling af, hvilke ting prøven skal teste og med hvilken vægtning. Ud fra denne matrix laves et *blueprint*, dvs. en oversigt over, hvilke stationer der anvendes til at teste hvilke områder. Den mest almindelige måde at etablere indholdsvaliditet på er at holde blueprintet op imod målbeskrivelsen for det kursus eller forløb, prøven skal teste.

I et blueprint vil man også kunne beskrive, hvilke kompetencer der testes i hvilket indhold. Eksempelvis vil man på en stand med anamneseoptagelse hos en patient med ondt i brystet teste anvendt viden om årsager til ondt i brystet og kompetence i at optage en anamnese. Her støder man ind i det problem, at man i flere studier har vist, at kompetencer er meget indholdsspecifikke [9-12]. Blueprinting er også et nyttigt redskab til efter prøven at undersøge, hvilke områder de studerende klarede sig dårligt i og dermed til at finde eventuelle svage punkter i undervisningen, curriculum eller prøven selv.

OSCE i denne form er blevet meget udbredt på sundhedsfaglige kandidatstudier verden over. Formatet kræver, at stationerne er reproducerbare, for at sikre at alle studenter møder det samme scenario. Det har medført brug af »standardiserede« patienter, dvs. personer, der er trænet i at fremstille en sygehistorie eller i nogle tilfælde at imitere objektive fund som lammelser og smertetilstande. Der anvendes også »simulerede patienter«, dvs. rigtige patienter med f.eks. hjertemislid, der præsenterer deres reelle kliniske tilstand i et simuleret scenario [11]. Den standardiserede patient kan anvendes som bedømmer og udfylder direkte efter interaktionen med den

Spørger studenten om:

	Ja	Nej
1. Lokalisation af bryst smerter?	<input type="checkbox"/>	<input type="checkbox"/>
2. Udstråling til hals/arm?	<input type="checkbox"/>	<input type="checkbox"/>
3. Varighed af bryst smerter?	<input type="checkbox"/>	<input type="checkbox"/>
4. Intensitet af bryst smerter?	<input type="checkbox"/>	<input type="checkbox"/>
5. Åndenød?	<input type="checkbox"/>	<input type="checkbox"/>
6. Tidligere bryst smerter?	<input type="checkbox"/>	<input type="checkbox"/>
7. Medicin taget i forbindelse med nuværende smerter?	<input type="checkbox"/>	<input type="checkbox"/>
8. Vanlig medicin?	<input type="checkbox"/>	<input type="checkbox"/>
9. Tidligere hjertesygdom?	<input type="checkbox"/>	<input type="checkbox"/>
10. Tidligere hypertension?	<input type="checkbox"/>	<input type="checkbox"/>
11. Sukkersyge?	<input type="checkbox"/>	<input type="checkbox"/>
12. Rygning?	<input type="checkbox"/>	<input type="checkbox"/>

Figur 1. Eksempel på tjekliste ved anamnese ved akutte bryst smerter.

studerende en tjekliste eller kommer med en udtalelse. Herunder kan mere bløde punkter som »patienttilfredshed« også medtages. Dette har speciel værdi, hvis man bruger OSCE i formativ hensigt, dvs. som en prøve, de studerende primært skal lære noget af. Den standardiserede patient kan give feedback til studenten ud fra konkrete punkter, og der er udviklet instrumenter til dette formål. Der er som regel ikke feedback efter den enkelte station i OSCE-prøver, der foretages i summativ hensigt, selv om de studerende formentlig, hvis tiden tillader det, vil være i stand til at modtage en sådan feedback [13].

Fra starten har OSCE været meget brugt til at teste procedurer med, og opfindsomheden har været stor. Hvis man ønsker at bedømme mere invasive procedurer, som f.eks. gynækologisk undersøgelse og lumbalpunktur, anvendes der fantomer, hvor den studerende skal demonstrere undersøgelsen for en eksaminator. Computere, video og simulatorer kan være anvendelige, hvis det er upraktisk eller umuligt at bruge levende personer. Eksempler på dette kan gå fra simple videooptagelser af epileptiske anfald til højteknologiske simulatorer som f.eks. avancerede genoplivningsmannequiner.

Efterhånden er OSCE blevet udviklet til at man ud over interpersonelle færdigheder også kan teste intellektuelle færdigheder som klinisk problemløsning. Det kan man gøre på forskellige måder. Man kan f.eks. bygge en case op ved i på hinanden følgende stationer at teste f.eks. anamneseoptagelse, objektiv undersøgelse, fortolkning af prøvesvar, stille diagnose og afgive forslag til behandling, alt eller dele heraf inden for den samme case. Hvor mange stationer, hver case skal omfatte, er en afvejning af, hvor mange kliniske områder man vil teste. En samling af sådanne cases, helst i alt 20-30 stationer, kan dække meget brede lægelige færdigheder. Den slags OSCE-prøver bliver derfor brugt også i postgraduat lægeuddannelse til autorisationsgivende eksaminer. Eksempler på lande, hvor man har indført sådanne, er Canada, USA og Australien.

I Danmark foregår den medicinske fagprøve for tredjelandslæger (læger, der kommer fra lande uden for EU) på den samme måde.

VIDENSKAB OG PRAKSIS | OVERSIGTSARTIKEL

OSCE har gennem årene gennemgået mange og uafhængige prøver af gennemførlighed, validitet og reliabilitet. Metoden er omfangsrigt dokumenteret i den medicinsk-pædagogiske litteratur og har opnået høj accept både hos lærere og medicinstuderende.

Effekter af OSCE

Al eksamination har en stærk styrende effekt på, hvad der bliver lært og hvordan. Hvis de studerende kun bliver eksamineret i teoretisk viden, lærer de de praktiske færdigheder dårligt. Allerede i 1981 kunne *Newble et al* vise, at hvis man indførte OSCE-prøve, forbedrede de studerende deres niveau af klinisk praktiske og intellektuelle færdigheder, uden at resultaterne af prøver i teoretisk viden blev forringet [14]. I de seneste år har man i to uafhængige undersøgelser fra Danmark vist, at de studerende selv bedømmer deres kliniske færdigheder som utilstrækkelige ved afslutning af medicinstudiet [15], og at deres aftagere er af den samme opfattelse [16]. Der findes således god grund til at indføre en eller anden form for OSCE-prøve i den kliniske fase af det danske medicinstudium.

Praktiske synspunkter

OSCE kræver et større antal medvirkende end de fleste andre prøveformer, men skal færdigheder testes på en objektiv måde, er der ingen bedre vej udenom. En prøve med f.eks. 20 stationer kræver en leder og en person (læge) pr. stand, der kan være eksaminator/bedømmer. På flere af stationerne kan yngre læger med fordel fungere som bedømmere. Hvis det er en skriftlig stand, kan lægen erstattes af teknisk personale. Herudover kræves der 1-2 personer på »gulvet« til teknisk assistance. Varigheden af prøven afhænger af antallet af stationer og stationslængden, men bør holdes under fire timer for den enkelte. Fyrrer studerende kan blive prøvet på en og samme dag. Ved store lægeskoler kan man oprette tidsmæssigt parallelle, men identiske prøver, hvilket fordobler kapaciteten. I Danmark er der tradition for, at den enkelte student-eksaminator-interaktion overvåges af en censor. I udlandet er dette ikke udbredt. Ofte er det tilstrækkeligt med en censor for hele prøven. Denne kan i forvejen bedømme prøvens kvalitet ved at gennemlæse casebeskrivelser, afkrydsningskemaer, opgaver mv. og ved selve prøven kun overvåge dens gennemførelse.

En person skal være ansvarlig for prøven. Tidsforbruget afhænger af opgavestillernes erfaring med prøveformen. Begynd minimum 3 mdr. før! Med tiltagende rutine bliver opgaven lettere.

En sidegevinst ved at lave en OSCE er den ofte interessante diskussion om standarder og forventninger, der opstår i processen og kan være givende for undervisningsplanlægningen efterfølgende.

Perspektiver

Skal danske medicinstuderendes kliniske færdigheder forbed-

res, er det uundgåeligt, at OSCE bliver brugt på nogle steder i studiet. I København forudsætter Studieplan 2000 en OSCE-prøve i slutningen af niende semester. Ved Syddansk Universitet er der siden 2002 blevet anvendt OSCE som prøve i obstetrik og pædiatri, samt planlagt en OSCE ved afslutningen af medicinstudiet.

Korrespondance: *Anne Lindebo Holm*, Kløvervænget 20 A 703, DK-5000 Odense C. E-mail: aholm@health.sdu.dk

Antaget: 16. april 2004

Interessekonflikter: Ingen angivet

Litteratur

1. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA* 2002;287:226-35.
2. O'Neill PA, Metcalfe D, David TJ. The core content of the undergraduate curriculum in Manchester. *Med Educ* 1999;33:121-9.
3. Rethans JJ, Norcini JJ, Baron-Maldonado M et al. The relationship between competence and performance: implications for assessing practice performance. *Med Educ* 2002;36:901-9.
4. Turnbull J, MacFadyen J, van Barneveld C et al. Clinical work sampling. *J Gen Intern Med* 2000;15:556-61.
5. Norcini JJ, Blank LL, Duffy FD et al. The mini-CEX: a method for assessing clinical skills. *Ann Intern Med* 2003;138:476-81.
6. Cuschieri A, Gleeson FA, Harden RM et al. A new approach to a final examination in surgery. *Ann R Coll Surg Engl* 1979;61:400-5.
7. Harden RM, Gleeson FA. Assessment of clinical competence using an objective structured clinical examination (OSCE). *Med Educ* 1979;13:41-54.
8. Hodges B, McNaughton N, Regehr G et al. The challenge of creating new OSCE measures to capture the characteristics of expertise. *Med Educ* 2002;36:742-8.
9. Fox R, Dacre J, McLure C. The impact of formal instruction in clinical examination skills on medical student performance – the example of peripheral nervous system examination. *Med Educ* 2001;35:371-3.
10. Hodges B, Turnbull J, Cohen R et al. Evaluating communication skills in the OSCE format: reliability and generalizability. *Med Educ* 1996;30:38-43.
11. Adamo G. Simulated and standardized patients in OSCEs: achievements and challenges 1992-2003. *Med Teach* 2003;25:262-70.
12. Roberts J, Norman G. Reliability and learning from the objective structured clinical examination. *Med Educ* 1990;24:219-23.
13. Hodder RV, Rivington RN, Calcutt LE et al. The effectiveness of immediate feedback during the objective structured clinical examination. *Med Educ* 1989;23:184-8.
14. Newble DI, Jaeger K. The effect of assessments and examinations on the learning of medical students. *Med Educ* 1983;17:165-71.
15. Mørcke AM, Eika B. De forventede praktiske færdigheder og de nyuddannede læger. *Ugeskr Læger* 2002;164:3531-4.
16. Ringsted CV, Tronnes H, Falck G. Lægers vurdering af deres kliniske færdighedsniveau ved påbegyndelse af turnus i foråret 1998. *Ugeskr Læger* 1999;161:5793-9.