

# Klinisk kompetencevurdering er central i alle speciallægeuddannelser

Martin G. Tolsgaard<sup>1,2</sup>, Doris Østergaard<sup>3,4</sup>, Lars Konge<sup>1</sup>, Torben V. Schroeder<sup>1</sup> & Charlotte Ringsted<sup>5,6</sup>

## STATUSARTIKEL

- 1) Center for Klinisk Uddannelse, Center for HR, Region Hovedstaden og Københavns Universitet  
 2) Juliane Marie Centret, Rigshospitalet  
 3) Dansk Institut for Medicinsk Simulation, Region Hovedstaden  
 4) Det Sundhedsvidenskabelige Institut, Københavns Universitet  
 5) Wilson Centret, University of Toronto  
 6) University Health Network, Canada

Ugeskr Læger  
 2014;176:V02140127

Klinisk kompetencevurdering (*workplace-based assessment*) dækker over testning af kompetence i virkelige situationer. Selvom kompetencevurdering er et relativt nyt begreb inden for mange kliniske specialer i Danmark, bygger det på lange traditioner fra Nordamerika. Her har testudvikling vundet tiltagende indpas siden 1. verdenskrig, hvor psykometri har forsynet militær og uddannelsesinstitutioner med begreber som pålidelighed (*reliability*) og validitet. Pålidelighed er et udtryk for et givent måleredskabs nøjagtighed. I klassisk testteori angives pålidelighed som den andel af en testscore, der kan tilskrives den »sande score« i forhold til tilfældig støj (fra f.eks. case, bedømmer og måleredskab) [1]. Validitet dækker over, hvorvidt testresultaterne er udtryk for den kompetence, der ønskes vurderet. Som eksempel vil multiple choice-questions (MCQ) kunne reflektere forskellige grader af teoretisk viden, men vil næppe være et validt mål for vurdering af praktiske kirurgiske færdigheder. Kompetencevurdering bør således tilpasses det, der ønskes vurderet.

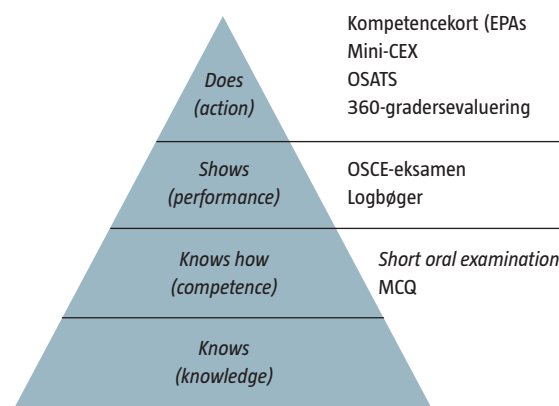
Millers pyramide er ofte anvendt til at illustrere, hvordan kompetencevurdering bør afstemmes med forskellige niveauer eller udtryk for kompetence – fra *knows* over *knows how*, *shows how* til *does* (Figur 1) [2]. Der findes mange forskellige kompetencevurderingsmetoder, der dækker over generelle såvel som specifikke redskaber. Således kan man bruge en Objective Structured Assessment of Technical Skills (OSATS) til specifik vurdering af kirurgiske færdigheder, en mini *clinical evaluation exercise* (CEX) til vurdering af håndteringen af en patientkonsultation eller en 360-gradersevaluering til en generel vurdering af en uddannelsessøgendes samarbejde med forskellige faggrupper [3].

Kompetencevurdering deles ofte op i summativ og formativ vurdering, afhængigt af formål med testen. Summativ vurdering er ofte synonymt med en testsituation, hvor formålet er at bedømme, om det opnåede kompetenceniveau er tilstrækkeligt til, at eksaminanden kan bestå. Formativ vurdering handler derimod om at give feedback til den uddannelsessøgende efter endt kompetencevurdering med henblik på at forbedre fremtidige præstationer. Grænserne er dog flydende, idet meningsfuld kompetencevurdering kræver, at niveauet er klart for den uddannelsessøgende og bedømmeren, samtidig med at enhver bedømmelse bør ledsages af feedback og instruktion [4]. I Tabel 1 gives der et overblik over hyppigt anvendte metoder.

Ønsket om pålidelige og valide test har i Nordamerika været drevet af juridiske såvel som patient-sikkerhedsmæssige hensyn og omfatter alle dele af uddannelsessystemet fra prægraduat uddannelse til certificering og recertificering på speciallægeniveau. Sundhedsstyrelsen har siden 2003 anbefalet, at klinisk kompetencevurdering bør indgå som bærende led i fremtidens speciallægeuddannelser [5], og udgav i 2013 en oversigt over tilgængelige kompetencevurderingsmetoder [6]. Da speciallægeanerkendelse i Danmark opnås ved godkendelse af obligatoriske uddannelseselementer og ikke gennem en »specialisteksamen«, er det danske system således baseret på løbende kompetencevurdering for at sikre, at de uddannelsessøgende har tilegnet sig den nødvendige kompetence, der forventes på speciallægeniveau.

FIGUR 1

Psykologen George Miller foreslog i 1990 en struktur for klinisk kompetencevurdering. Eksempler på kompetencevurderingsmetoder, der kan bruges inden for hvert niveau, er anført ud for hver kategori.



CEX = *clinical evaluation exercise*; EPAs = *entrusted professional activities*; MCQ = *multiple choice-questionnaire*; OSATS = *Objective Structured Assessment of Technical Skills*; OSCE = *Objective Structured Clinical Examination*.

## KOMPETENCEVURDERINGS ANVENDELIGHED

For at sikre størst mulig anvendelighed af kompetencevurderingsmetoderne bør deres *pålidelighed*, *validitet*, *omkostninger*, *brugeraccept* og *læringsværdi* tages i

betragtning [7]. De anvendte metoder bør have en *pålidelighed*, der sikrer nøjagtige målinger og tilstrækkelig evidens for deres *validitet* for at sikre, at testresultaterne afspejler den kompetence, der vurde-



TABEL 1

Oversigt over hyppigt anvendte kompetencevurderingsmetoder.

Metode/anvendelsesområde	Indhold	Fordele/værd at vide
<b>Mini-CEX</b> Vurdering af kompetence i varetagelsen af patientkonsultationen	<i>9-punktsskalaer med vurdering af</i> Anamneseoptagelse Objektiv undersøgelse Empati og professionel adfærd Klinisk vurdering Patientvejledning Effektivitet og organisationsevne Overordnet klinisk kompetence	<i>Fordele</i> Giver struktur til feedback og er vidt udbredt  <i>Værd at vide</i> Dårlig prædiktiv validitet og beskeden udvikling i scorer over tid
<b>OSATS</b> Vurdering af kirurgisk kompetence	<i>5-punktsskalaer med vurdering af</i> Respekt for vævet Tid og bevægelser Metodehåndtering Kendskab til metoder Anvendelsen af assistent Fremdrift i operation og planlægning Kendskab til den specifikke procedure	<i>Fordele</i> Overskuelig og med god validitetsevidens inden for flere typer kirurgi  <i>Værd at vide</i> Supplerende procedurespecifik tjekliste kan være gavnlig for at give meningsfuld feedback
<b>360-gradersevaluering</b> Vurdering af professionel adfærd, kliniske kompetence og samarbejde	Selvevaluering samt vurderinger foretaget af omkring 10 kolleger inden for samme og andre specialer Der findes ingen konsensus om indhold., men spørgsmålene omfatter ofte evnen til samarbejde, kommunikation og evt. kognitive elementer	<i>Fordele</i> Kan under optimale forhold fremme samarbejde og tillid  <i>Værd at vide</i> Subjektive vurderinger fra selvvalgte bedømmere medfører oftest overvejende positive bedømmelser Væsentlig risiko for bedømmerbias
<b>CBD</b> Struktureret diskussion af case med kliniske vejledere	Gennemgang af caserelevante aspekter <i>Emner til vurdering</i> Definition af klinisk problemstilling Gennemgang af tilgængelig information Prioritering af muligheder Konsekvenser af handlinger Retfærdiggørelse af beslutning Ethiske og professionel adfærd Samarbejde med andre faggrupper	<i>Fordele</i> Muliggør vurdering og dokumentation af daglige kliniske casediskussioner  <i>Værd at vide</i> Kræver over 6 vurderinger for at sikre tilstrækkelig pålidelighed, såfremt CBP skal anvendes til summative vurderinger Potentiel bias i selektion af cases
<b>NOTSS</b> Redskab til vurdering af ikke-tekniske færdigheder	<i>Kan anvendes til under et operativt indgreb at vurdere kirurgens ikke-tekniske færdigheder inden for disse kategorier</i> Situationsbevidsthed Beslutningstagning Kommunikation Ledelse	<i>Fordele</i> Anvendes til at strukturere observation, vurdere og give feedback til en uddannelsessøgende læge på dennes ikke-tekniske færdigheder efter en operation  <i>Værd at vide</i> Der findes manual tilgængelig på <a href="http://www.regionh.dk/dims">www.regionh.dk/dims</a>
<b>Kompetencekort</b> Kompetencekort til vurdering af specifikke arbejdsopgaver eller færdigheder	Et kompetencekort kan indeholde beskrivelse af specifikke færdigheder, som den uddannelsessøgende skal mestre selvstændigt Disse kompetencer varierer fra fag til fag, men kan f.eks. omfatte stuegang i intern medicin, kopanlæggelse i obstetrik, brug af tvang i psykiatrien eller procedurer inden for anæstesi	<i>Fordele</i> Trækker på flere lægeroller og har relevante konsekvenser for patientsikkerhed og beslutninger om tildeling af ansvar i klinikken  <i>Værd at vide</i> Kræver kobling af uddannelsesprogrammets mål med de specifikke opgaver, den uddannelsessøgende tildeles i det pågældende speciale

CBP = casebaseret diskussion; CEX = *clinical evaluation exercise*; MCQ = multiple choice-questionnaire; NOTSS = Non-Technical Skills for Surgeons; OSATS = Objective Structured Assessment of Technical Skills.

res. Der er i den medicinske uddannelseslitteratur bred enighed om, at kompetence bør vurderes gentagne gange og med forskellige metoder for at sikre en høj pålidelighed og validitet i kompetencevurderingen. Desuden anbefales det at bruge flere forskellige bedømmere for at mindske bias i vurderingerne. Bias kan dog aldrig helt undgås.

Nyere forskning peger i retning af, at globale vurderinger, der trækker på vejlederes overordnede vurdering af den uddannelsessøgende, formentlig er mere valide udtryk for kompetence end summen af mange detaljerede vurderinger [8]. Et dansk studie inden for anæstesiologisk uddannelse viste imidlertid, at klinikerne foretrak tjeklister frem for globale vurderinger, hvilket forfatterne tilskrev, at tjeklister tydeliggjorde uddannelsesprogrammets læringsmål for de kliniske vejledere og de uddannelsessøgende [9]. De økonomiske og tidsmæssige omkostninger, der er forbundet med kompetencevurdering, er afhængige af den valgte metode, og man bør derfor anvende redskaber, hvormed man kan tage udgangspunkt i arbejdssituationer, og som kan fungere i en travl klinisk hverdag. De kliniske vejlederes og uddannelsessøgendes accept af kompetencevurdering har stor betydning ikke kun for pålideligheden og validiteten af testresultaterne, men også for implementering af kompetencevurdering som et fast element på de enkelte afdelinger. Kvalificering af de kliniske vejledere spiller derfor en central rolle, men også omfanget af kompetencevurdering er vigtig.

Et dansk studie viste, at kompetencevurderingsmetoderne hurtigt reduceredes til avancerede afkrydsningsøvelser, hvis de kliniske vejledere ikke opfattede redskaberne som autentiske i testsituationen [10]. Den sidste faktor, som er af betydning for metoderne anvendelighed, er deres læringsværdi. I et større review har man påvist sparsom evidens for effekten af kompetencevurdering på lægers læring [11]. Dette review omfatter dog primært generiske kompetencevurderingsredskaber såsom mini-CEX,

360-gradersevaluering og direkte observation af procedurer. Kun i ét tilfælde (360-gradersevalueringen) er der dokumenteret positiv effekt på læring, men her blev kompetencevurderingen ledsaget af vejledende mentorsamtaler, hvilket muligvis har haft indflydelse på resultaterne.

I et nyligt publiceret hollandsk studie undersøgte man de uddannelsessøgendes og de kliniske vejlederes oplevelser og erfaringer med kompetencevurdering [12]. De fleste uddannelsessøgende værdsatte muligheden for struktureret feedback, men mange mente samtidig, at kompetencevurdering var en ekstra arbejdsbyrde, hvis udbytte afhæng af organisationens evne til at inkorporere den som en del af produktionen i klinikken. I et nyere dansk studie har man ligeledes påvist, at hvis nye uddannelsesinnovationer skal blive en succes, skal der være støtte fra alle niveauer i uddannelsessystemet, og at incitamentet til kompetencevurdering derfor ikke alene bør påhvile de uddannelsessøgende [13].

Hvis kompetencevurderingen skal medføre positiv effekt på læring, må man derfor sikre grundig oplæring af kliniske vejledere. I de tilfælde, hvor kompetencevurderingen er formativ, skal den følges af meningsfuld feedback, idet en kompetencevurderingsscore i sig selv har begrænset læringsværdi [14]. I simulationsforskningen findes modeller for, hvordan feedback kan bruges til at afdække, hvilke overvejelser og følelser samt hvilken viden der ligger til grund for den uddannelsessøgendes handlinger [15]. Disse modeller giver mulighed for at stimulere til refleksion hos den uddannelsessøgende og kan hjælpe de kliniske vejledere til at formidle kritik på en konstruktiv måde [16].

## KOMPETENCEVURDERING OG PATIENTSIKKERHED

Kompetencevurdering har ikke blot betydning for de uddannelsessøgendes læring, hensynet til patientsikkerhed spiller også en rolle. Set fra et patientsikkerhedsperspektiv har kompetencevurdering til formål at identificere uddannelsessøgende, der ikke lever op til bestemte standarder. *Birkmeyer et al* foretog et studie af OSATS-bedømmelser i forbindelse med bariatrisk kirurgi og påviste, at kirurgernes OSATS-scoring var associeret til hyppigheden af operative såvel som perioperative komplikationer [17]. I et større britisk studie blev kompetencevurderinger fra en population af 1.646 uddannelsessøgende læger analyseret med henblik på at undersøge, om disse scorer kunne bruges til identificering af uddannelsessøgende, der havde vanskeligheder. Den prædiktive værdi af kompetencevurderingerne var dog meget lav, og den gruppe af læger, der kom i vanskeligheder, scorede ikke signifikant forskelligt fra de øvrige uddannelses-



### FAKTABOKS

Sundhedsstyrelsen anbefaler, at klinisk kompetencevurdering bør indgå som bærende led i speciallægeuddannelserne.

Kompetencevurderingsmetoder anvendelighed er afhængig af deres pålidelighed, validitet, omkostninger, brugeraccept og læringsværdi.

Succesfuld implementering af kompetencevurdering kan kun ske ved kritisk valg af kompetencevurderingsmetoder til eget speciale, uddannelse af vejledere i at anvende disse og ledelsesopbakning til at sikre, at kompetencevurdering kan foretages i en travl klinisk hverdag.

Brugen af kompetencekort giver mulighed for, at tildeling af ansvar baseres direkte på den uddannelsessøgendes kompetencer.

søgende på halvdelen af de undersøgte kompetencevurderingsmetoder [18].

Der er således evidens for, at kompetencevurdering kan reflektere adfærd, der har betydning for patientsikkerhed, men at kompetencevurderingsscorerne næppe kan bruges til forudsigelse af, hvordan de uddannelsessøgende vil præstere i fremtiden.

## KOMPETENCEKORT

I speciallægeuddannelsen er et af formålene med kompetencevurdering at bedømme den uddannelsessøgende med henblik på at give vedkommende større ansvars- og opgaveområder. I den sammenhæng kan tolkningen af de mere generiske kompetencevurderinger vise sig at være problematisk. Dels bliver de uddannelsessøgende ofte vurderet urealistisk højt af deres kliniske vejledere, og dels kan der opstå tvivl om fortolkningen af testresultatet blandt de kliniske vejledere til trods for, at der imellem dem hersker enighed om de enkelte uddannelsessøgendes præstationer [19].

Begreberne betroede professionelle aktiviteter (*entrusted professional activities* (EPAs)) og milepæle har derfor vundet tiltagende indpas. Til bedømmelse af EPAs anvendes kompetencekort, der kan opfattes som et »kørekort«, som tildeles den uddannelsessøgende, når denne har udvist evne til selvstændigt at udføre en specifik arbejdsopgave på tilfredsstillende vis. Kompetencekort sigter mod holistisk vurdering af hele opgaver, der er relevante for den uddannelsessøgendes kliniske virke [20] og kan således involvere flere lægeroller og mange delkompetencer – f.eks. i forbindelse med stuegang på en internmedicinsk afdeling [21]. Milepæle omfatter udviklingsmæssige beskrivelser af, hvorledes og hvornår den uddannelsessøgende kan forventes at flytte sig på vej mod tiltagende klinisk selvstændighed [22].

I Danmark er f.eks. godkendt introduktionsuddannelse eller oprykning i vagtlag at betragte som milepæle. Hvor tolkningen og konsekvenserne af en OSATS, en mini-CEX eller en casebaseret diskussion kan være svært gennemskuelige, frembyder kompetencekort og milepæle en kliniknær form for kompetencevurdering, der omfatter den holistiske bedømmelse og respekterer de kliniske vejlederes globale dømmekraft. Inden for anæstesiologisk uddannelse har Danmark været blandt de første i verden til at etablere strukturerede forløb med brug af kompetencekort, EPAs og milepæle – f.eks. med hvilke kompetencekort der skal være bestået inden for en bestemt tidsramme, eller hvilke der skal være bestået, for at man kan gå i vagt [23].

På Dansk Selskab for Anæstesi og Intensiv Medicins hjemmeside findes der eksempler på kompeten-

cekort til vurdering af tekniske og ikke-tekniske færdigheder.

## UDFORDRINGER OG FREMTIDENS KOMPETENCEVURDERING I DANMARK

Kompetencevurdering har en central plads inden for lægelig videreuddannelse, men en succesfuld implementering kan kun ske ved kritisk valg af kompetencevurderingsmetoder til eget speciale, uddannelse af vejledere i at anvende disse og ledelsesopbakning til at sikre, at kompetencevurdering kan foretages i en travl klinisk hverdag. En OSATS-vurdering og en 360-gradersevaluering uden veldefinerede formål og uddannelsesmæssige konsekvenser bliver nemt synonymt med avancerede afkrydsningsøvelser – og i så fald kan kompetencevurdering forekomme meningsløs for såvel de uddannelsessøgende som de kliniske vejledere [12].

Kompetencevurdering kræver grundig oplæring af brugerne, tid i det kliniske arbejde og meningsfulde konsekvenser af vurderingerne i form af klart definerede kriterier for, hvilke funktioner den uddannelsessøgende må varetage selvstændigt [24]. Brugen af kompetencekort er ét godt bud på, hvordan disse udfordringer kan imødekommes og integreres på en meningsfuld måde i fremtidig klinisk praksis, således at tildeling af ansvar baseres direkte på den uddannelsessøgendes kompetencer.

## SUMMARY

Martin G. Tolsgaard, Doris Østergaard, Lars Konge, Torben V. Schroeder & Charlotte Ringsted:

The assessment of clinical skills is imperative in postgraduate specialty training

Ugeskr Læger 2014;176:V02140127

According to the Danish Ministry of Health recommendations, workplace-based assessment (WPBA) should be a central component in postgraduate specialty training. In this review the potential impact of WPBA with regard to patient safety, medical licensing and effects on doctors' learning is explored. The utility of WPBA instruments is reviewed with respect to their reliability, validity, cost, acceptability, and educational impact. Finally, the concept of Entrusted Professional Activities is discussed in view of the utility framework, and recommendations for future practice are provided.

**KORRESPONDANCE:** Martin G. Tolsgaard, Center for Klinisk Uddannelse, Center for HR, Region Hovedstaden, Blegdamsvej 9, 2100 København Ø.  
E-mail: martintolsgaard@gmail.com

**ANTAGET:** 3. juni 2014

**PUBLICERET PÅ UGESKRIFTET.DK:** 15. september 2014

**INTERESSEKONFLIKTER:** Forfatterens ICMJE-formularer er tilgængelige sammen med artiklen på Ugeskriftet.dk

## LITTERATUR

1. van der Vleuten CP, Schuwirth LW. Assessing professional competence: from methods to programmes. *Med Educ* 2005;39:309-17.
2. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;65(suppl 9):S63-7.
3. Downing SM, Yudkowsky R. Assessment in health professions education. New York: Taylor & Francis, 2009.
4. Watling C, Driessen E, van der Vleuten CP et al. Beyond individualism: professional culture and its influence on feedback. *Med Educ* 2013;47:585-94.
5. Sundhedsstyrelsen 2003. [www.sst.dk](http://www.sst.dk) (9. jul 2014).
6. Kompetencevurderingsmetoder, en oversigt. Sundhedsstyrelsen, 2013. <http://sundhedsstyrelsen.dk/publ/Publ2013/08aug/KompetenceVurdMetoder.pdf> (9.jul 2014).
7. van der Vleuten C. The assessment of professional competence: developments, research and practical implications. *Adv Health Sci Educ Theory Pract* 1996;1:41-67.
8. Hodges B. Assessment in the post-psychometric era: learning to love the subjective and collective. *Med Teach* 2013;35:564-8.
9. Ringsted C, Østergaard D, Ravn L et al. A feasibility study comparing checklists and global rating forms to assess resident performance in clinical skills. *Med Teach* 2003;25:654-8.
10. Tolsgaard MG, Arendrup H, Lindhardt BO et al. Construct validity of the reporter-interpreter-manager-educator structure for assessing students' patient encounter skills. *Acad Med* 2012;87:799-806.
11. Miller A, Archer J. Impact of workplace based assessment on doctors' education and performance: a systematic review. *BMJ* 2010;341:c5064.
12. Fokkema JP, Teunissen PW, Westerman M et al. Exploration of perceived effects of innovations in postgraduate medical education. *Med Educ* 2013;47:271-81.
13. Tolsgaard MG, Arendrup H, Pedersen P et al. Feasibility of self-directed learning in clerkships. *Med Teach* 2013;35:1409-15.
14. Bsted KM, Malling B, Skjelsager K et al. Bedømmerbias i speciallægeuddannelsen. *Ugeskr Læger* 2011;73:2788-90.
15. Steinwachs B: How to facilitate a debrief. *Simul Gaming* 1992;23:186-95.
16. Rudolph JW, Simon R, Dufresne RL et al. There's no such thing as "nonjudgmental" debriefing: a theory and method for debriefing with good judgment. *Simul Healthc* 2006;1:49-55.
17. Birkmeyer JD, Finks JF, O'Reilly A et al. Surgical skill and complication rates after bariatric surgery. *N Engl J Med* 2013;369:1434-42.
18. Mitchell C, Bhat S, Herbert A et al. Workplace-based assessments of junior doctors: do scores predict training difficulties? *Med Educ* 2011;45:1190-8.
19. Ginsburg S. Respecting the expertise of clinician assessors: construct alignment is one good answer. *Med Educ* 2011;45:546-8.
20. ten Cate O, Scheele F. Competency-based postgraduate training: can we bridge the gap between theory and clinical practice? *Acad Med* 2007;82:542-7.
21. Nørgaard K, Ringsted C, Dolmans D. Validation of a checklist to assess ward round performance in internal medicine. *Med Educ* 2004;38:700-7.
22. Hicks PJ, Englander R, Schumacher DJ et al. Pediatrics milestone project: next steps toward meaningful outcomes assessment. *J Grad Med Educ* 2010;2:577-84.
23. Ringsted C, Østergaard D, van der Vleuten CP. Implementation of a formal in-training assessment programme in anaesthesiology and preliminary results of acceptability. *Acta Anaesthesiol Scand* 2003;47:1196-203.
24. Ringsted C, Henriksen AH, Skaarup AM et al. Educational impact of in-training assessment (ITA) in postgraduate medical education: a qualitative study of an ITA programme in actual practice. *Med Educ* 2004;38:767-77.