

Objective structured clinical examination (OSCE) in der medizinischen Ausbildung: Eine Alternative zur Klausur

J.-F. Chenot, M. Ehrhardt¹

Zusammenfassung

Es wird angenommen, dass mündliche und schriftliche Prüfungen von Medizinstudenten eher Faktenwissen als klinische Kompetenz beurteilen. Eine Alternative stellt die Objective structured clinical examination (OSCE) dar, ein Parcours von Stationen, an denen praktische Fähigkeiten beurteilt werden. Die Übersicht fasst internationale Erfahrungen mit OSCE zusammen, unter besonderer Berücksichtigung von Validität, Reliabilität und Objektivität. Die Testgütekriterien können erreicht werden, erfordern jedoch eine sorgfältige Planung und Durchführung.

Schlüsselwörter

Objective structured clinical examination, klinische Kompetenz, Validität, Reliabilität, Objektivität.

Summary

Objective structured clinical examination (OSCE) in medical training: An alternative to a written exam

It is assumed that oral and written exams of medical students assess knowledge rather than clinical competency. An alternative approach is the objective structured clinical examination (OSCE), where students rotate round a series of stations and clinical skills are assessed. The review summarizes international experience with OSCE with regard to validity, reliability and objectivity. Those measurements of test quality can be achieved but require substantial efforts in design and implementation.

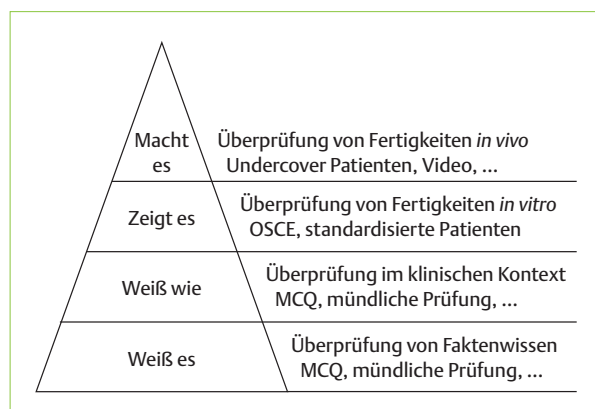
Key words

Objective structured clinical examination, clinical competency, validity, reliability, objectivity

Die bisher üblichen Prüfungsformen

Die verbreitetste Prüfungsform in der medizinischen Ausbildung in Deutschland ist die schriftliche Prüfung. Sie basiert meistens auf Multiple Choice-Fragen (*multiple choice questions* [MCQ]), daneben werden noch mündliche Prüfungen durchgeführt. Da Prüfungsformate erhebliche Auswirkungen auf das Lernverhalten der

Studierenden haben (1), sind in Deutschland gute theoretische Kenntnisse zunächst nützlicher als praktische Fertigkeiten. Es bestehen jedoch Zweifel, ob die bisher verwandten Prüfungsformen geeignet sind, klinische Kompetenz zu messen (Abb. 1) (2, 3).



■ Abbildung 1: System zur Beurteilung von Fertigkeiten nach Miller

Alternative Prüfungsformen

Im Rahmen der neuen Approbationsordnung müssen scheinpflichtige Fächer ab Oktober 2003 benotet werden. Dies bedeutet eine Aufwertung der universitären Veranstaltungen und das Ende der so genannten »Sitzscheine«.

Diese Änderungen können genutzt werden, um alternative Prüfungsformen für die klinische Kompetenz, wie die Objective Structured Clinical Examination (OSCE) als alleinige Prüfung oder als Teil einer zusammengesetzten Prüfung zu etablieren.

Dr. med. Jean-François Chenot, MPH

Abteilung Allgemeinmedizin, Universität Göttingen

Humboldtallee 38, 37073 Göttingen

E-Mail: jchenot@gwdg.de

¹ Institut für Allgemeinmedizin Hamburg

Systematische Literatursuche zur OSCE

Suchmethode: Es wurde eine systematische Literatursuche in MEDLINE für den Zeitraum Januar 1975 bis Mai 2003 durchgeführt. Zusätzlich wurden die Referenzen der gefundenen Artikel gescreent und eine ergänzende Handsuche in den einschlägigen Fachzeitschriften durchgeführt. Folgende Schlagwörter (MeSH) wurden verwendet:

- *objective structured clinical examination* (435 Artikel)
- OSCE (326 Artikel)

Alternative Begriffe, mit denen ebenfalls eine Literatursuche durchgeführt wurde, sind *clinical performance assessment*, *clinical skill assessment* und *standardized patients*.

Studienauswahl: Berücksichtigt wurden Veröffentlichungen, die sich auf das Studium der Humanmedizin bezogen. Artikel, die sich mit der OSCE in der Weiterbildung beschäftigten, wurden nur berücksichtigt, sofern sie sich mit wichtigen methodischen Fragen auseinandersetzen.

Was ist eine OSCE?

OSCE wurde 1975 erstmals von Harden vorgestellt (5). Eine allgemeingültige Definition gibt es nicht. Grundstruktur ist ein Parcours mit Stationen, wie z.B. die Erhebung einer Anamnese oder eine körperliche Untersuchung. Gewöhnlich schließt sie einige Stationen mit standardisierten Patienten (SP) ein. An anderen Stationen wird mit Modellen, z.B. Reanimationspuppen (6), Videos (7) oder Computern (8) gearbeitet. Es können auch Stationen mit MCQs eingebaut werden. Eine OSCE hat 10 bis 25 Stationen, die zwischen fünf und zehn Minuten dauern. Studierende werden an den Stationen von Prüfern nach standardisierten Kriterien bewertet.

Diese Prüfung soll nicht wie bisher in mündlichen oder schriftlichen Prüfungen Wissen, sondern Fertigkeiten und klinische Kompetenzen prüfen. Sie soll auch objektiver sein als die bisher zur praktischen Prüfung übliche Untersuchung eines einzelnen echten Patienten (*»long case«*). Das Verhalten von Patienten und ihre Mitwirkung z.B. bei der Anamneseerhebung ist Schwankungen unterworfen und die Variationen der zu Verfügung stehenden Patienten (*intercase-reliability*) erschwert die Objektivierbarkeit (9). OSCE versucht durch Standardisierung die Variabilität zwischen verschiedenen Patienten und Prüfern zu überwinden.

Erfahrungen mit der OSCE bestehen in Deutschland unter anderem an den Universitäten in Ulm (10), Münster, Hannover (11), Göttingen, Witten/Herdecke (12) und

Düsseldorf. Das OSCE-Format wird häufig in der Notfallmedizin eingesetzt, ohne dass explizit von einer OSCE gesprochen wird.

Welche Fächer und Teilgebiete können mit OSCE geprüft werden?

Das flexible OSCE-Format wird in vielen Fächern, Teilgebieten und auf allen Ausbildungsstufen im Medizinstudium eingesetzt (Tabellen 1 und 2).

Tabelle 1: Fachgebiete, in denen die OSCE zur Evaluation von Studierenden eingesetzt wurde

Fachgebiet	Literatur
Allgemeinmedizin	21, 72, 73
Chirurgie	74–77
Geriatric	78
Gynäkologie	79–81
Notfallmedizin	82–85
Pädiatrie	86–88
Psychiatrie	89, 90
Radiologie	91
Rheumatologie	92
Urologie	93

Tabelle 2: Teilgebiete, in denen die OSCE zur Evaluation von Studierenden eingesetzt wurde

Teilgebiet	Literatur
Suchtmedizin	94, 95
Digitale rektale Untersuchung	96
Ernährungsmedizin	97
Ethik in der Medizin	98–100
Evidenzbasierte Medizin	101–103
Körperliche Untersuchung	104–106
Kommunikation	50, 107–109
Raucheranamnese	110
Schmerztherapie bei Krebs	111
Arzneimittelverordnung	112
Umweltmedizin	113

Prinzipiell kann zwischen Lehr- und Prüfungs-OSCEs unterschieden werden. Neben der Abschlussevaluation ist die Validierung neuer Unterrichtskonzepte oder der Vergleich mit traditionellen Lehrmethoden ein weiteres Einsatzgebiet (15, 16).

In fast allen Publikationen wird optimistisch und positiv über Erfahrungen sowohl von Studierendenseite als auch der Fakultät berichtet. Studierende halten die OSCE für fair (19) und ziehen sie den MCQ vor, bevorzugen aber die Untersuchung eines echten Patienten (13). Dies wird mit Problemen der Studierenden beim Übergang zur OSCE begründet.

Diese Unsicherheit legt sich mit zunehmenden Erfahrungen mit der OSCE (20). Studierende stimmen auch zu, dass die OSCE-Ergebnisse klinische Kompetenz reflektieren, zweifeln aber, ob diese die Vermittlung der Fähigkeiten in der Ausbildung reflektiert (21). Studierende berichten über größeres Selbstvertrauen in ihre praktischen Fertigkeiten nach einer klinischen Basisfähigkeiten-OSCE (22).

Nur vereinzelt finden sich negative Äußerungen zur OSCE. Studierende waren unzufrieden mit der Überprüfung ihrer kommunikativen Fähigkeiten und hielten diese prinzipiell für nur eingeschränkt prüfbar (23). Ein Prüfer empfand die Teilnahme als eine langweilige Zumutung (24). Es lässt sich ein erheblicher Publikationsbias vermuten, da es wahrscheinlich nach dem großen Aufwand für die Ein- und Durchführung einer OSCE nur selten zu negativen Publikationen kommt.

Wie valide, reliabel und objektiv ist OSCE?

Prüfungen werden nach drei Testgütekriterien beurteilt (25):

- Validität
- Reliabilität
- Objektivität

Validität, Reliabilität, Objektivität

Die **Validität** gibt an, wie gut der Test in der Lage ist, genau das zu messen was er vorgibt zu messen.

Die **Reliabilität** kennzeichnet den Grad der Genauigkeit, mit dem das geprüfte Merkmal gemessen wird.

Die **Objektivität** gibt an, in welchem Ausmaß die Testergebnisse vom Testanwender unabhängig sind.

Validität

Die Validität ist das wichtigste Testgütekriterium. Die OSCE hat eine hohe Augenscheinvalidität (Inhaltsvalidität), wie die bereits oben erwähnte positive Einschätzung der meisten Berichte zeigt. Diese kann nicht numerisch bestimmt werden und beruht allein auf subjektiven Einschätzungen.

Da es keinen Goldstandard zur Identifikation von nicht ausreichend qualifizierten Studierenden gibt, kann OSCE nur gegen einen anderen Test verglichen werden. Dies wird numerisch durch einen Korrelationskoeffizienten (*concurrent validity*) ausgedrückt (Werte zwischen 0,4 und 0,6 gelten als mittelmäßig, Werte über 0,6 als hoch). Eine andere Möglichkeit, die Validität zu beschreiben, ist

die Vorhersage der Ergebnisse in einem anderen Test, z.B. im Abschlussexamen (*predictive validity*).

Es zeigt sich meist eine schlechte bis mittlere Korrelation zwischen den Prüfungsformen (Tabelle 3). Dies spiegelt unter anderem wider, dass die OSCE andere Parameter als MCQ-Examen erfasst. Klinische Erfahrung sollte sich in klinischer Kompetenz ausdrücken. Vergleiche von Probanden auf verschiedenen Ausbildungsstufen können zur Beurteilung der Validität und der Reliabilität eingesetzt werden. Das dies nicht immer der Fall ist, konnte sowohl für klassische mündliche und schriftliche Prüfungen als auch für die OSCE gezeigt werden (2, 27). Wird das in einer OSCE gezeigte Verhalten auch in die tägliche Praxis umgesetzt? In einer der wenigen vorhandenen Studien bei niederländischen Hausärzten wurden die Ergebnisse einer OSCE und einer schriftlichen Prüfung mit Videoaufnahmen in der Sprechstunde verglichen. Es zeigte sich nur eine mittlere Korrelation, sowohl für die OSCE als auch die schriftliche Prüfung (28). In einer anderen Studie korrelierte die von standardisierten Patienten beurteilte Leistung von Rheumatologen in der Sprechstunde gut mit den OSCE-Ergebnissen (29).

Reliabilität

Die Fähigkeit einer Prüfung, kompetente von nicht-kompetenten Studierenden zu unterscheiden (Sensitivität, Spezifität), ist ein Aspekt der Reliabilität. Entscheidender Faktor hierbei ist die Festsetzung der Bestehensgrenze (30, 31, 32). Möglichkeiten sind das Setzen eines absoluten Grenzwertes durch den Autor der OSCE-Station oder eines relativen durch die schwächsten Prüflinge (*borderline group method [BGM]*). Die letztere Methode wird bevorzugt und erzeugt meist niedrigere Grenzwerte als das Festsetzen einer absoluten Bestehensgrenze (33). Eine weitere Alternative ist die so genannte Angoff-Methode, die sich als einfach durchführbar und reliabel erwiesen hat (32, 34). In einer Gruppendiskussion einigen sich Experten durch Schätzung auf einen absoluten Grenzwert (35). Im Vergleich aller Studien wird die BGM bevorzugt (14, 36).

Ein anderer Aspekt der Reliabilität, der auch die Auswertungsobjektivität berührt, ist die Konsistenz der Messung zwischen verschiedenen Prüfern, d.h., kommen verschiedene Prüfer zu einem vergleichbaren Ergebnis? Werte für diese Interraterreliabilität ab 0,8 werden als gut bezeichnet. Die in den Studien berichtete Interraterreliabilität schwankt erheblich zwischen 0,11 und 0,97 (37–40).

Es wird davon ausgegangen, dass die Varianz in der einzelnen Stationsbeurteilung sich mit zunehmender Länge

Tabelle 3: Vergleich der OSCE mit anderen Prüfungsformen bei Studierenden (a: USMLE 1 (United States Licensing Examination 1) entspricht etwa dem 1. Staatsexamen, b: SP standardisierte Patienten, c: EMQ Extended matching Question, d: SAQ Short answer Questions, e: NBME (National Board of Medical Examiners) entspricht etwa dem IMPP)

Autor	Land	N	OSCE	Alternative Evaluation	Korrelation mit der OSCE	Kommentar
Simon et al. (114)	USA	355	Interdisziplinäre OSCE 16 Stationen	USMLE 1 ^a MCQ	Korrelationskoeffizient 0,41 (p<0,001)	Mittlere prädiktive Validität. Stationen ohne SPb korrelierten besser mit USMLE-Ergebnis.
Remmen et al. (115)	NL	47 59	Klinische Basisfähigkeiten 12 Stationen, Checklisten & Globalbeurteilung	Klinische Basisfähigkeiten MCQ	Korrelationskoeffizient Checkliste 0,41–0,48 Globalbeurteilung 0,35–0,41	MCQ mäßig valide Alternative zur Evaluation klinischer Basisfähigkeiten
Van der Vleuten et al. (116)	NL	380	Klinische Basisfähigkeiten 8 Stationen Checklisten	Klinische Basisfähigkeiten MCQ	Korrelationskoeffizient 0,72	Schriftlicher Test der OSCE vergleichbar.
Wilkinson et al. (14)	NZ	181 188 203	18 Stationen 3 Zentren über 3 Jahre 2 unabhängige Prüfer Checklisten & Globalbeurteilung	Abschlusssexamen MCQ Kursevaluation	0,6–0,63 (p<0,01) 0,62–0,67 (p<0,01)	Hohe Konstruktvalidität und Reliabilität
Wass et al. (117)	UK	79 214	Interdisziplinäre OSCE Abschlusssexamen 20 Stationen 2 unabhängige Prüfer	Abschlusssexamen »long case« 2 nicht standardisierte (echte) Patienten 2 unabhängige Prüfer	Ergebnisse werden als projizierte Korrelation angegeben. Erst nach 10 »long cases« wird eine Korrelation von 0,84–0,88 erreicht	»Long case« ist der OSCE ebenbürtig
Wass et al. (71)	UK	214	interdisziplinäre OSCE Abschlusssexamen 20 Stationen (2 h 20 min) plus 2 »long cases«	Abschlusssexamen MCQ (3 h) EMQc (20 min) SAQd (3 h)	Korrelationskoeffizient 0,28 0,77 0,78	Identische Studierenden-Gruppe wie in der vorhergehenden Studie. Reliabilität Gewichtungenabhängig zwischen 0,77 und 0,83
Prislin et al. (21)	USA	335	Allgemeinmedizin-OSCE 22–27 Stationen über 4 Jahre	MCQ (NBMEe) Kursevaluation	Korrelationskoeffizient 0,222 (p<0,001) 0,262 (p<0,001)	Geringe Korrelation, hohe Konstruktvalidität.
Townsend et al. (72)	UAE	28	Allgemeinmedizin-OSCE (Prä- und Posttest) 10 Stationen Checklisten & Globalbeurteilung	Abschlusssexamen Kursevaluation	Korrelationskoeffizient 0,337 (p=0,08) 0,784 (p=0,0005)	Mäßige Korrelation mit Abschlusssexamen, gute mit Kursevaluation
Collins et al. (13)	NZ	608	Abschlusssexamen interdisziplinäre OSCE 28 Stationen über 6 Jahre	Abschlusssexamen MCQ »long case«	Korrelationskoeffizient 0,51 (p<0,001) 0,33 (p<0,001)	Nur Kombination der Prüfungsformen erlaubt valide Beurteilung.
Matsell et al. (87)	Ca	77	Pädiatrie-OSCE mit 10 Stationen	MCQ Praktikumevaluation Gesamtnote	Korrelationskoeffizient 0,65 (p<0,001) 0,54 (p<0,001) 0,79 (p<0,001)	Gute Validität.
Martin et al. (31)	NZ	266	interdisziplinäre OSCE im ersten klinischen Jahr 23 Stationen	Andere Examen, OSCEs und MCQs in den folgenden Jahren	Studierende, die im ersten Jahr durch den OSCE gefallen sind, sind 6-mal häufiger bei einem späteren Examen durchgefallen.	Innovative Darstellung der Ergebnisse als ROC-Kurve. Gute prädiktive Validität.

der OSCE ausgleicht (117). Die Zahl der Stationen spielt auch für die Reliabilität eine Rolle. Die Variabilität der Leistung eines Prüflings an verschiedenen Stationen (interstation reliability) wird als Reliabilitätskoeffizient (Alphakoeffizient von Cronbach) ausgedrückt. Ab einem

Wert von 0.8 geht man von einer guten Reliabilität aus. Mit zunehmender Zahl der Stationen steigt der Reliabilitätskoeffizient. Es werden mindestens zehn Stationen als notwendig angesehen (41). Die berichteten Reliabilitätsmaße reichen von schlecht bis gut (Tabelle 4).

Tabelle 4: Übersicht über die Reliabilität der OSCE

Literatur	Land	N	OSCE	Reliabilitätsmaß	Kommentar
Hodges et al. (118)	Ca	60 Assistenten 30 Studierende	Kommunikations-OSCE mit 18 Stationen	Interrater-Reliabilität 59–0,63 Cronbach-Alpha 0,17 (Checkliste) 0,2 (Globale Beurteilung)	Schlechte Reliabilität
Singer et al. (98)	USA	88 Studierende	Ethik-OSCE 4 standardisierte Patienten Video der Prüfung	Reliabilitätskoeffizient 0,28 Interrater-Reliabilität -0,8 Cronbach-Alpha 0,07 IntraKlassen-Korrelationskoeffizient 0,94	Geringe Reliabilität, sollte mit anderen Prüfungsformen kombiniert werden Gute Reliabilität
Humphris et al. (7)	UK	200 Studierende	Kommunikations-OSCE (OSVE) Video der Prüfung		
Regehr et al. (44)	Ca Ca	53 Assistenten 53 Assistenten	Basisfertigkeiten-OSCE	Cronbach-Alpha 0,79 (Checkliste) 0,85 (Globale Beurteilung) 0,89 (Kombination)	Gute Reliabilität
Matsell et al. (87)	Ca	77 Studierende	Pädiatrie-OSCE mit 10 Stationen	Variation des Cronbach-Alpha bei 4 Durchgängen 0,12–0,69	Gute Reliabilität
Wilkinson et al. (49)	NZ	204 Studierende	Anamnese-OSCE mit 11 Stationen	Cronbach-Alpha 0,65 Interrater-Reliabilität 74	Gute Reliabilität

Objektivität

Durch Standardisierung einer Prüfung soll eine möglichst hohe Objektivität und damit Gerechtigkeit und Fairness garantiert werden. Aspekte der Bewertungsobjektivität wie die Interraterreabilität wurden zum Teil bereits behandelt. Andere Aspekte sind der Einfluss des Evaluationsmodus, des Prüfers, der Prüfreihefolge und des Prüflings selbst auf die Ergebnisse.

Es gibt zwei prinzipielle Möglichkeiten der Bewertung einer OSCE-Station: Entweder globale Beurteilungen oder die traditionell eingesetzten Checklisten (42, 43). Diese beiden Methoden zeigen eine variable Korrelation und Reliabilität (Tabelle 5). Checklisten sind inhaltspezifischer, während die globale Beurteilung ein breiteres Spektrum an Fertigkeiten erfasst (44).

Die Mutmaßung der Prüflinge über die Evaluationsmethode hatte Einfluss auf das Prüfungsergebnis. Studierende, die annahmen, nur mit einer Checkliste beurteilt zu werden, zeigten konsistentere Leistungen (45). In einer OSCE, in der Allgemeinärzte und Assistenten verglichen wurden, zeigten Assistenten bei der Beurteilung mit Checkliste bessere und in der globalen Beurteilung schlechtere Ergebnisse (46). Kliniker mit einem höheren Kompetenzlevel lösen Probleme offensichtlich mit anderen Strategien, als sie von Checklisten vorgegeben werden (47). Checklisten können zu einem ritualisierten Verhalten im Umgang mit den standardisierten Patienten (SP) führen. Eine Kombination der beiden Bewertungsstrategien kann die Reliabilität erhöhen (48, 72). Die Bewertung kann bei Stationen mit SP durch diese selbst, durch externe Bewerter oder durch beide erfol-

gen. Es zeigt sich eine gute Übereinstimmung zwischen SP und Ärzten bzw. dem Gesamtscore (49), wobei Ärzte technische Fertigkeiten kritischer bewerten (50). Keinen Einfluss auf die Bewertung hat die Erfahrung des Prüfers, der Ausbildungsstand (Assistenten versus Fachärzte), die Fachrichtung (Allgemeinärzte versus Fachärzte anderer Richtungen) (39, 50), das Geschlecht des Prüfers oder des Prüflings (56, 57), sowie die Ermüdung der Prüfer im Verlauf der OSCE (50, 51).

Auch Informationen, die später geprüfte Studierende von ihren bereits geprüften Kommilitonen über OSCE-Stationen erhalten haben, hatten keine signifikanten Auswirkungen auf die Leistungen (52). Es gibt keine Hinweise darauf, dass ausländische Studierende oder Angehörige ethnischer Minderheiten durch OSCE benachteiligt sind (53–55). Die OSCE-Ergebnisse korrelieren kaum mit dem Vorbereitungsaufwand der Studierenden, sondern eher mit deren allgemeinen Leistungen (58).

Wie werden OSCE-Ergebnisse mitgeteilt?

Es wurden keine Berichte gefunden, wie die OSCE-Ergebnisse mitgeteilt werden. Prinzipiell können Ergebnisse dichotom als »Bestanden« oder »Durchgefallen« oder auf einer Skala zurückgemeldet werden. Aus solchen Rückmeldungen erhält der Studierende nur wenige Informationen über seine Schwächen. Eine Kombination von Notenvergabe und zusätzlichem individuellem Feedback wird zum Beispiel in Australien und Südafrika angewandt (59). In einem primär als Unterrichtsmittel eingesetzten OSCE erhielten Studierende nach jeder

Tabelle 5: Vergleich der OSCE mit MCQ und der mündlichen Prüfung

	Mündliche Prüfung (119)	MCQ	OSCE
Stufe der klinischen Kompetenz nach Miller	Faktenwissen	Faktenwissen	Fertigkeiten z.B.: • Kommunikation • Körperliche Untersuchung
Vorbereitungsaufwand	gering	hoch	hoch
Personalaufwand	hoch	gering	hoch
Zeitaufwand Prüfer	abhängig von der Zahl der Prüfer	relativ gering	hoher Vorbereitungsaufwand, hoher Prüferbedarf, lange Prüfungszeiten am Stück
Objektivität	Subjektivität der Prüfer, gleicht sich nur bei mehreren Prüfern aus	anonymisiert objektiv	Subjektivität der Prüfer gleicht sich aus
Validität	variabel	geringe Augenscheinvalidität in Bezug auf klinische Kompetenz	hohe Augenscheinvalidität
Reliabilität	variabel, abhängig von der Testdauer	gut bis sehr gut	variabel, abhängig von der Anzahl der Stationen
Standardisierbarkeit	nur eingeschränkt standardisierbar	standardisiert	standardisierbar Problem: Interrater-Reliabilität

Kommunikation. Eine Kombination mit der Seminar- oder anderen Beurteilungen, z.B. Evaluation des Blockpraktikums zu einer Gesamtnote ist möglich. Eine faire Gewichtung der OSCE-Leistung in einer Gesamtnote hängt jedoch entscheidend von der Kenntnis der Reliabilität ab (71). Unabhängig von der Prüfungsform wird eine Standardisierung der allgemeinenmedizinischen Lehre notwendig, denn geprüft und benotet werden kann nur, was vermittelt worden ist. Dies stellt eine große Herausforderung für die Allgemeinmedizin mit ihren vielen extrauniversitären Lehrärzten dar. OSCE bietet eine hervorragende Möglichkeit, Schwächen im

Station direkt Rückmeldung vom Prüfer, was sehr positiv gewertet wurde (17).

Curriculum und bei den Lehrmethoden zu identifizieren und kann so zur Verbesserung der Ausbildung beitragen.

Wie hoch ist der finanzielle Aufwand für eine OSCE?

Die Erfahrungen aus dem Ausland lassen sich aufgrund unterschiedlicher personeller und organisatorischer Strukturen nicht einfach auf Deutschland übertragen (60–63). Der finanzielle Aufwand für eine OSCE mit 20 Stationen wurde in Hannover Mitte der 90er Jahre mit Fixkosten von 8.000–10.000,- DM bzw. ca. 130,- DM pro Student angegeben (64). Damit ist eine OSCE sicherlich teurer als eine MCQ-Prüfung oder eine mündliche Prüfung, zudem wurde der Aufwand für die Planung und Pilotierung in der Kalkulation nicht berücksichtigt. Trotz der höheren Kosten einer OSCE überwiegt jedoch die Meinung, dass sich der finanzielle Aufwand lohnt (65).

OSCE in der Allgemeinmedizin

Bisher wird OSCE in der Allgemeinmedizin im Ausland vornehmlich in der Weiterbildung eingesetzt (66–69), so zum Beispiel als Teil der Facharztprüfung in Kanada. OSCE eignet sich gut zur Prüfung der Kernelemente des Fachs, wie klinischen Basisfähigkeiten oder Arzt-Patient-

Zusammenfassung

OSCE ist eine Methode zur Prüfung klinischer Kompetenz mit hoher Augenscheinvalidität. Sie wurde zur Prüfung in vielen Fächern erfolgreich als alleinige oder ergänzende Prüfung eingesetzt. Ein Vergleich von mündlicher Prüfung, MCQ-Prüfung und OSCE ist in *Tabelle 5* zusammengefasst. OSCE ist kein einheitlich standardisiertes Prüfungsformat, so dass die Ergebnisse einzelner Studien nicht verallgemeinert werden können. Die Übersicht zeigt, dass die Testgütekriterien Validität, Reliabilität und Objektivität erfüllt werden können, sie müssen aber für jede OSCE neu etabliert werden. Sie sind abhängig von der Sorgfalt der Vorbereitung und Durchführung der Prüfung.

Danksagung: Für das kritische Lektorat und die Korrektur bedanken wir uns bei A. Kapmeyer, B. Reichel, Dr. T. Fischer und Prof. M.M. Kochen.

Literatur

Literatur bei den Verfassern

Nützliche Hinweise für die Planung einer OSCE

Internet

- www.meds.queensu.ca/~ohse/OSCE.html (Queen's University Kingston, Ontario)
Hinweise zur praktischen Durchführung einer OSCE
- www.kmrrec.org/KM/osce/ (Kessler Medical & Rehabilitation, New Jersey)
Umfangreiches Material, Patientenevaluationsbögen, Prüfungsbögen, Zeitpläne und eine OSCE Musterstation (Knie)
- <http://www.healthsystem.virginia.edu/internet/cstap/> (University of Virginia)
Praktische Hinweise zum Umgang mit Standardisierten Patienten
- <http://www.utmed.com/physdx/osce.htm> (Houston Medical School)
Beispiel für einen OSCE-Abschlussevaluationsbogen, Anleitung für Studierende

Literatur

- Selby C, Osman L, Davis M, Lee M: How To Do It: Set up and run an objective structured clinical exam. BMJ 1995; 310: 1187–90
- Dornan T, O'Neill P: Core clinical skills for OSCEs in medicine. Churchill Livingstone Edinburgh 2000
Eine OSCE-Anleitung für Studenten mit vielen Anregungen zur Gestaltung von OSCE-Stationen

Zur Person



Dr. med. Jean-François Chenot, MPH, Jahrgang 1969, amerikanischer Facharzt Primary Care/Internal Medicine, Gesundheitswissenschaftler, Arzt in Weiterbildung zum Facharzt für Allgemeinmedizin in einer Landarztpraxis, wissenschaftlicher Mitarbeiter der Abteilung Allgemeinmedizin der Universität Göttingen seit September 2001. Forschungsschwerpunkte: »Kreuzschmerzen«, Integration von EbM in die Praxis und die studentische Lehre.