

OSCE の評価者の違いによる評価の一致性に関する検討

岩 堀 正 俊¹⁾ 小 川 雅 之¹⁾ 広 瀬 俊¹⁾
米 田 博 紀¹⁾ 住 友 伸一郎²⁾ 村 松 泰 徳²⁾
大 森 俊 和¹⁾ 都 尾 元 宣¹⁾ 藤 原 周¹⁾

Study on Conformity between Difference of Inspectors of the OSCE Results

IWAHORI MASATOSHI¹⁾, OGAWA MASAYUKI¹⁾, HIROSE SHUN¹⁾, YONEDA HIRONORI¹⁾, SUMITOMO SHINICHIROU²⁾,
MURAMATSU YASUNORI²⁾, OHMORI TOSHIKAZU¹⁾, MIYAO MOTONOBU¹⁾ and FUJIWARA SHUU¹⁾

平成17年12月より「臨床実習開始前の共用試験」が正式に導入された。特に総括的評価として実施されるOSCEではその信頼性、妥当性が担保されている必要があり、歯学分野でもその点に関しての検討がみられるようになってきた。しかし実際には各評価項目の点数は評価者の主観に影響を受けることも多いために、評価のばらつきによる客観性の低下が懸念される。

今回 OSCE において評価される、精神・運動領域、認知領域、情意領域の各領域の特性、各評価者による評価のばらつきを検討するために一致率、カッパ係数を求めた。

その結果、評価する領域別の評価者ペアの評価の一致の度合いは、認知領域、情意領域、精神運動領域の順に高かった。

評価者ペア間一致性が著しく低い場合、評価する項目や評価の方法が適切でない場合がある。

同一の課題において他のペアと比べて評価の一致が著しく低い評価者ペアが存在した。

キーワード：OSCE，評価，一致率，カッパ係数

OSCE is a method, introduced by Harden of objectively evaluating clinical ability and also is a method of objectively evaluating clinical efficacy of the technique and the attitude, etc. that cannot be measured by descriptive examination. It has been applied as a method of evaluating the clinical abilities of medical students and residents in Europe and America because the results can be given as feedback to students and inspectors. In Japan, the Common Achievement Test before clinical training was formally introduced in December, 2005. It is necessary to show the reliability and validity of OSCE as an overall evaluation. Research into this area has been initiated in the dental field; however, the evaluation is often influenced by inspectors subjectivity; decreased objectivity because of differences in the evaluation is feared. External OSCE inspectors have been obligated to complete an inspector training workshop since 2006, because it is demanded that the conformity of inspector evaluations is high.

In this research, conformity and kappa the coefficient were calculated to examine the characteristics of the mental-exercise, recognition and affect areas for differences in inspector evaluations.

The average conformity in the recognition area between inspector pairs showed the highest score, followed by the affect area and then the mental-exercise area.

Key words: OSCE, Evaluation, Conformity, Kappa coefficient

¹⁾朝日大学歯学部口腔機能修復学講座歯科補綴学分野
²⁾朝日大学歯学部口腔病態医療学講座口腔外科学分野
501 0296 岐阜県瑞穂市穂積1851

¹⁾Department of Prosthodontics, Division of Oral Functional Science and Rehabilitation

²⁾Department of Oral and Maxillofacial Surgery, Division of Oral Pathogenesis and Disease Control
Asahi University School of Dentistry
Hozumi 1851, Mizuho, Gifu 501 0296, Japan
(平成20年11月28日受理)

緒 言

OSCE は、Harden¹⁾らによって紹介された臨床能力を客観的に評価する方法であり、記述試験では測定できない技術や態度などの臨床能力を客観的に評価する方法である。その結果を学習者および教育者にフィードバックが可能ことから欧米各国では医学生やレジデントの臨床能力を評価する方法として応用されてきた²⁻⁴⁾。我が国においても平成17年12月より「臨床実習開始前の共用試験」が正式実施として導入された⁵⁾。特に総括的評価として実施される OSCE ではその信頼性、妥当性が担保されている必要があり、歯学分野でもその点に關しての検討がみられるようになってきた。しかし実際には各評価項目の点数は評価者の主観に影響を受けることも多いために、評価のばらつきによる客観性の低下が懸念される。平成18年度からの「臨床実習開始前の共用試験」では外部評価者は評価者養成ワークショップ修了者であることが義務づけられている⁵⁾。これは、上で述べたように複数評価者の評価の一致度が高いことが要求されているからである。

今回 OSCE において評価される、精神・運動領域、認知領域、情意領域の各領域の特性、各評価者による評価のばらつきを検討するために一致率、カッパ係数を求め考察を行った。

対象および方法

1. OSCE の概要

2003年度の朝日大学附属病院歯科医師臨床研修医のプログラムは、4月から9月を前期、10月から3月を後期とした。後期研修では10月より3ヶ月間、研修医の選択により協力型施設（15人）および附属病院（12人）で研修を行った。前期研修修了時に前期 OSCE（2003年7月21日実施）を、後期研修修了時に後期 OSCE（2004年3月18日実施）を行った。課題の概要を表1、2に示した。課題数は前期 OSCE5課題、後期 OSCE5課題の合計10課題であった。

2. OSCE 実施方法

評価方法は1名の臨床研修医に対して2名の評価者が各項目について2段階および3段階の評価を行い評価シートに記入した。OSCE の評価者は前日に各課題責任者より評価マニュアルを基本としてキャリブレーションを行った。前期 OSCE の評価者13名、後期 OSCE の評価者40名であった。

3. 一致性の検討

前期および後期 OSCE を併せたすべての課題の評価項目は、田口らの方法⁸⁾に従い、情意領域、精神運動領域、認知領域の3領域に区分された。すなわち、

表1 前期 OSCE の課題と GIO

課題1. 印象採得 (GIO) 考究用模型作製のための印象採得の技能および態度を習得する。
課題2. 医療面接（急性症状） (GIO) 急性症状を訴える初診患者のとの良好な人間関係の構築と問題点の抽出を行うために、医療面接についての基本的な態度、技能および知識を習得する。
課題3. 単純抜歯器具の選択 (GIO) 単純抜歯を行うために必要な器具に関する知識と技能を習得する。
課題4. 刷掃指導 (GIO) 軽度歯周炎患者に対しての刷掃指導を行うための基本的な態度、技能および知識を習得する。
課題5. 10枚法エックス線写真のマウント (GIO) 正確に口内法エックス線写真フィルムがマウントできる知識を習得する。

表2 後期 OSCE の課題と GIO

課題1. 医療面接（急性症状） (GIO) 急性症状を訴える初診患者のとの良好な人間関係の構築と問題点の抽出を行うために、医療面接についての基本的な態度、技能および知識を習得する。
課題2. ラバーダム防湿 (GIO) 安全かつ効率的な歯科治療を行うために、ラバーダム防湿についての基本的な体位度、技能および知識を習得する。
課題3. 支台歯形成 (GIO) 下顎左側第一大臼歯に全部鋳造冠の遠心隣接面の支台歯形成を行う技能を習得する。
課題4. バイタルサイン (GIO) 患者と良好なコミュニケーションをとりながら、適切にバイタルサインを得るための知識、態度、技能を修得する。
課題5. 刷掃指導 (GIO) 軽度歯周炎患者に対しての刷掃指導を行うための基本的な態度、技能および知識を習得する。

情意領域では前期 OSCE11項目、後期 OSCE18項目の合計29項目、精神運動領域では前期 OSCE11項目、後期 OSCE32項目、合計43項目、認知領域では前期 OSCE17項目、後期 OSCE25項目、合計42項目である。

1 受験者につき評価を行った2名の評価者の評価の一致性について偶然性を加味した検討を行うためにすべての評価ペアおよび評価項目でカッパ係数および一致率を算出した。また、情意、精神運動および認知領域で評価者による評価の一致性に差が生じるかを検討する目的で、前期・後期のカッパ計数値を合計し、領域ごとのカッパ係数の平均値の差を対応のないt検定で検定を行った。

統計処理ソフトはEXCEL 統計 Ver5 α (エスミ社製)を使用した。各評価者を区別するために任意な英文字を1字から3字をつけた。なお、前期 OSCE の課題3「単純抜歯器具の選択」および課題5「10枚法エックス線写真のマウント」は全ての評価者間で全ての評価が一致したので統計処理からは除外した。これらの2課題は全ての評価項目が認知領域であった。

結 果

表3に前期 OSCE の、表4に後期 OSCE の、情意領域の評価項目における全ての評価者ペアのカッパ係数および一致率を示した。以下は同様に表5、6は精神運動領域の、表7、8は認知領域の結果である。

表3および表4に示した情意領域における統計処理

表3 前期 OSCE における情意領域の評価者間の一致率とカッパ係数

課題1 印象採得								
評価者HM	VS	評価者FS	評価者OT	VS	評価者IM	評価者YK	VS	評価者YT
評価段階	3×3		評価段階	3×3		評価段階	3×3	
一致率	0.93		一致率	0.87		一致率	1.00	
カッパ係数	0.86		カッパ係数	0.75		カッパ係数	1.00	
課題2 医療面接								
評価者YAK	VS	評価者OT	評価者OT	VS	評価者YK	評価者KK	VS	評価者SS
評価段階	3×3		評価段階	3×3		評価段階	3×3	
一致率	0.63		一致率	0.80		一致率	0.75	
カッパ係数	0.32		カッパ係数	0.65		カッパ係数	0.56	
課題4 刷掃指導								
評価者KS	VS	評価者ST	評価者NM	VS	評価者SI			
評価段階	2×2		評価段階	2×2				
一致率	0.94		一致率	0.77				
カッパ係数	0.84		カッパ係数	0.50				

表4 後期 OSCE における情意領域の評価者間の一致率とカッパ係数

課題1 医療面接								
評価者OT	VS	評価者KM	評価者MK	VS	評価者MM	評価者SM	VS	評価者FM
評価段階	3×3		評価段階	3×3		評価段階	3×3	
一致率	0.73		一致率	0.94		一致率	0.90	
カッパ係数	0.55		カッパ係数	0.90		カッパ係数	0.86	
課題2 ラバダム防湿								
評価者YJ	VS	評価者SI	評価者HM	VS	評価者HN	評価者YAK	VS	評価者NA
評価段階	2×2		評価段階	2×2		評価段階	2×2	
一致率	0.95		一致率	0.86		一致率	0.90	
カッパ係数	0.90		カッパ係数	0.72		カッパ係数	0.79	
課題3 文台書形成								
評価者MS	VS	評価者OA	評価者OK	VS	評価者YO	評価者KT	VS	評価者NT
評価段階	2×2		評価段階	2×2		評価段階	2×2	
一致率	0.93		一致率	0.86		一致率	0.64	
カッパ係数	0.84		カッパ係数	0.63		カッパ係数	0.29	
課題4 バイタルサイン								
評価者TA	VS	評価者TT	評価者KK	VS	評価者HM	評価者SM	VS	評価者SS
評価段階	3×3		評価段階	3×3		評価段階	3×3	
一致率	0.93		一致率	0.86		一致率	0.79	
カッパ係数	0.86		カッパ係数	0.77		カッパ係数	0.32	
課題あ5 刷掃指導								
評価者KK	VS	評価者HM	評価者KS	VS	評価者ST	評価者AJ	VS	評価者SM
評価段階	2×2		評価段階	2×2		評価段階	2×2	
一致率	0.93		一致率	0.71		一致率	0.71	
カッパ係数	0.63		カッパ係数	0.24		カッパ係数	0.00	

表5 前期 OSCE における精神運動領域の評価者間の一致率とカッパ係数

課題1 印象採得								
評価者HM	VS	評価者FS	評価者OT	VS	評価者IM	評価者YK	VS	評価者YT
評価段階	3×3		評価段階	3×3		評価段階	3×3	
一致率	0.78		一致率	0.85		一致率	0.94	
カッパ係数	0.63		カッパ係数	0.76		カッパ係数	0.90	

表 6 後期 OSCE における精神運動領域の評価者間の一致率とカッパ係数

課題1 医療面接									
評価者OT	VS	評価者KM	評価者MK	VS	評価者MM	評価者SM	VS	評価者FM	
評価段階	3×3		評価段階	3×3		評価段階	3×3		
一致率	0.69		一致率	0.83		一致率	0.86		
カッパ係数	0.52		カッパ係数	0.73		カッパ係数	0.76		
課題2 ラバダム防護									
評価者YJ	VS	評価者SI	評価者HM	VS	評価者HN	評価者YAK	VS	評価者NA	
評価段階	2×2		評価段階	2×2		評価段階	2×2		
一致率	0.73		一致率	0.76		一致率	0.47		
カッパ係数	0.56		カッパ係数	0.60		カッパ係数	0.15		
課題3 支台歯形成									
評価者MS	VS	評価者OA	評価者OK	VS	評価者YO	評価者KT	VS	評価者NT	
評価段階	3×3		評価段階	3×3		評価段階	3×3		
一致率	0.93		一致率	0.71		一致率	0.57		
カッパ係数	0.55		カッパ係数	0.42		カッパ係数	0.11		
課題5 バイタルサイン									
評価者TA	VS	評価者TT	評価者KK	VS	評価者HM	評価者SM	VS	評価者SS	
評価段階	3×3		評価段階	3×3		評価段階	3×3		
一致率	0.91		一致率	0.83		一致率	0.89		
カッパ係数	0.76		カッパ係数	0.69		カッパ係数	0.72		

表 7 前期 OSCE における認知領域の評価者間の一致率とカッパ係数

課題1 印象採得									
評価者HM	VS	評価者FS	評価者OT	VS	評価者IM	評価者YK	VS	評価者YT	
評価段階	3×3		評価段階	3×3		評価段階	3×3		
一致率	1.00		一致率	0.80		一致率	1.00		
カッパ係数	1.00		カッパ係数	0.41		カッパ係数	1.00		
課題4 刷掃指導									
評価者KS	VS	評価者ST	評価者NM	VS	評価者SI				
評価段階	3×3		評価段階	3×3					
一致率	0.82		一致率	0.81					
カッパ係数	0.63		カッパ係数	0.62					

表 8 後期 OSCE における認知領域の評価者間の一致率とカッパ係数

課題2 ラバダム防護											
評価者YJ	VS	評価者SI	評価者HM	VS	評価者HN	評価者YAK	VS	評価者NA	評価者IM	VS	評価者YK
評価段階	2×2		評価段階	2×2		評価段階	2×2		評価段階	2×2	
一致率	0.86		一致率	0.93		一致率	0.82		一致率	0.92	
カッパ係数	0.69		カッパ係数	0.86		カッパ係数	0.60		カッパ係数	0.82	
課題3 支台歯形成											
評価者MS	VS	評価者OA	評価者OK	VS	評価者YO	評価者KT	VS	評価者NT	評価者YZ	VS	評価者IM
評価段階	3×3		評価段階	3×3		評価段階	3×3		評価段階	3×3	
一致率	1.00		一致率	1.00		一致率	0.71		一致率	0.96	
カッパ係数	1.00		カッパ係数	1.00		カッパ係数	0.55		カッパ係数	0.93	
課題4 バイタルサイン											
評価者TA	VS	評価者TT	評価者KK	VS	評価者HM	評価者SM	VS	評価者SS	評価者TS	VS	評価者OA
評価段階	3×3		評価段階	3×3		評価段階	3×3		評価段階	3×3	
一致率	0.86		一致率	0.86		一致率	0.86		一致率	1.00	
カッパ係数	0.68		カッパ係数	0.70		カッパ係数	0.78		カッパ係数	1.00	
課題5 刷掃指導											
評価者KK	VS	評価者HM	評価者K	VS	評価者ST	評価者AJ	VS	評価者SM	評価者KM	VS	評価者KM
評価段階	2×2		評価段階	2×2		評価段階	2×2		評価段階	2×2	
一致率	0.85		一致率	0.86		一致率	0.74		一致率	0.88	
カッパ係数	0.68		カッパ係数	0.70		カッパ係数	0.78		カッパ係数	1.00	

の対象となる評価ペア数の合計は28組であり、前後期併せた一致率の平均は0.84、カッパ係数は0.66であった。課題別で見ると前期 OSCE では課題1；印象採得の一致性は比較的高く、課題2；医療面接の一致性は低かった。後期 OSCE では課題2；ラバーダム防湿の一致性が高い傾向が見られた。また課題1の医療面接（評価者 NKvsSS：0.29）、課題3；支台歯形成（評価者 KTvsNT：0.29）、課題5；刷掃指導（評価者 KSvsST：0.24）と極めて低い値を示した。

表5および表6に示した精神運動領域における統計処理の対象となる評価ペア数の合計は19組であり、前後期併せた一致率の平均は0.78、カッパ係数は0.57であった。前期 OSCE では精神運動領域の評価項目を含んだ課題は課題1；印象採得だけであった。後期 OSCE では課題5；バイタルサインの一致性が比較的高く、課題3；支台歯形成でカッパ係数が各評価者間で0.55、0.42、0.11、0.35と低い一致性を示した。

表7および表8に示した認知領域における統計処理の対象となる評価ペア数の合計は21組であり、前後期併せた一致率の平均は0.87、カッパ係数は0.78であった。認知領域では、前・後期 OSCE を通じて、他の領域と比較して高いカッパ係数が得られる傾向が観察され低い項目でも、前期 OSCE の課題1の印象採得（評価者 OTvsIM：0.41）であった。

情意領域、精神運動領域および認知領域のカッパ係数の平均値の比較では、認知領域0.78、情意領域0.66、精神運動領域0.57の順に高い値を示し、情意・認知領域および精神運動領域・認知領域間において有意な差を認めた（表9・11）。

考 察

OSCE は従来の筆記試験では、評価できない精神運動領域（技能）および情意領域（態度）に係わる臨床能力を、シミュレーションテストや実地技能などを行うことで、客観的に評価できることから広く世界各国でも利用されている。本邦における歯科医学教育領域では平成17年12月より「臨床実習開始前の共用試験」が正式実施されている。このように、総括的評価として実施される OSCE は、特にその評価の信頼性および妥当性が重要である。OSCE に係る信頼性・妥当性に関する報告は、相沢ら⁹⁾ OSCE 評価の信頼性・妥当性を高めるには、評価項目における、評価領域が明瞭になる文章表現にすることおよび評価者間の事前確認で「何を評価するのか」を統一する必要があることを報告している。伊藤らは、複数の評価者の評価が一致しない原因としてヒューマンエラーを観点とした検討を行い、面接のコンテンツでは「聞き逃し」が多

表9 情意領域と精神運動領域のカッパ係数の比較

	情意領域	精神運動領域	差 (X1-X2)
件数	28.00	19.00	9.00
平均	0.66	0.57	0.09
標準偏差	0.20	0.21	-0.01
統計量	1.46		
自由度	37		
0.5%点	2.72		
2.5%点	2.03		
P値	0.15		
判定マーク	[]		

表10 情意領域と認知領域のカッパ係数の比較

	情意領域	認知領域	差 (X1-X2)
件数	28	21	7
平均	0.66	0.78	-0.12
標準偏差	0.20	0.18	0.02
統計量	2.27		
自由度	45.00		
0.5%点	2.69		
2.5%点	2.01		
P値	0.03		
判定マーク	[*]		(有意差あり)

表11 精神運動領域と認知領域のカッパ係数の比較

	精神運動領域	認知領域	差 (X1-X2)
件数	19	21	-2
平均	0.57	0.78	-0.21
標準偏差	0.21	0.18	0.03
統計量	3.42		
自由度	36.00		
0.5%点	2.72		
2.5%点	2.03		
P値	0.00		
判定マーク	[**]		(有意差あり)

く、プロセスでは「行動の解釈」や「判定基準の不一致」が多くみられると報告している¹⁰⁾。説明系 OSCE では開始直後の2人目までの受験者に対する評価がそれ以後の評価と比べて低くなることが報告されている¹¹⁾。

本研究では、評価者間の評価の一致性を領域別に検討をした。後期情意領域の医療面接で低いカッパ係数であった評価者ペア（NKvsSS）は同様に精神運動領域でも一致が著しく低かった。

精神運動領域では後期 OSCE 課題3の支台歯形成における各評価者間でカッパ係数が低かった。全ての評価者間で一致性が低かったため、精神運動領域の評価項目ごとにカッパ係数を算出した（表12）。3.支台歯形成後の評価、である「辺縁の位置」0.07、「辺縁の形態」-0.03、「隣接面の削除の有無」0.84、「軸面傾斜度の確認」0.09が特に低い値を示した。支台歯

表12 後期 OSCE 支台歯形成の評価シートとカップパ係数（精神運動領域）

評価項目および評価シート	カップパ係数
3. 支台歯形成中の注意	
1) 薬, 中指によってタービンを固定をした場合・・・1点 固定しなかった場合・・・0点	0.46
2) 注水下で削除した場合・・・1点 注水しなかった場合・・・0点	0.64
4. 支台歯形成後	
1) 辺縁の位置	
形成部全てが歯肉縁と形態, 高さが一致している場合・・・2点 部分的に乱れている場合・・・1点 部分的に一致している・・・0点	0.07
2) 辺縁形態	
形成部が全てシャンファー形態の場合・・・2点 の一部がシャンファー形態の場合・・・1点 がシャンファー形態以外の場合・・・0点	-0.03
3) 隣接歯の削除の有無	0.84
削除のない場合・・・2点 僅かでも削除した場合・・・0点	
4) 軸面傾斜度の確認	0.09
5度以内の場合(適切な軸面傾斜度)・・・2点 5～10度の場合(ほぼ適切な軸面傾斜度)・・・1点 10度以上の場合(不適切)・・・0点	

形成後に辺縁の位置や形態を OSCE のタイムスケジュールを守りながら短時間で、均一に評価するのは非常に困難であったためと考えられ、適切な評価がなされなかった可能性がある。事前にキャリブレーションを十分に行っても、あるいは評価マニュアルを工夫してもこのような評価項目の不安定な評価は避けがたいと考える。

一方、「隣接面の削除の有無」や 4. 支台歯形成中の注意の「注水下で削合した」は、隣接歯を確認することが容易であることや、実際に注水されているかは一目瞭然であり、判断に迷うことがない。にもかかわらず一部の評価者ペア（KTvsNT）は他の領域でも著しく低いカップパ係数を示した。このように別の評価者ペアと比較して著しくカップパ係数が低くなる評価者ペア、NKvsSS および YTvsNT には評価者自身、評価者の組み合わせ等に何らかの問題があると推測できる。特に、総括的な評価を行う OSCE の場合や、受験者に適当なフィードバックをするためには評価者を適正に評価して、不適切な評価をする評価者の評価能力を改善することが必要である。

領域別に一致の度合いを観察すると認知領域が運動精神領域および情意領域よりもペアごとのカップパ係数は高かった。相澤ら¹²も同様に認知領域が最も高く、情意領域が最も低いという結果を報告している。今回

検討した課題の中で認知領域に分類される「単純抜歯器具の選択」および「10枚法エックス線写真のマウント」では全ての評価者間で全ての評価が一致した。これは正しい器具選択やマウントを行えたかを評価者が答え合わせをするだけであり、他の項目で評価されるコンテットやプロセスが含まれてないので、評価が完全に一致したと考えられる。

本研究では、評価者の選定に関しては可能な限り課題と評価者の専門分野を一致させるようにした。1 評価者は 1 課題のみ評価を行っている。また、前日に評価の擦り合わせを行えるだけ評価にバラツキが出ないように留意した。しかし評価者の中でもワークショップ修了者と未修了者が混在していることもあり、なかなか完全に一致した評価を得るのは困難であった。今後ワークショップ修了者は増加していくので評価の一致性は高くなっていくのではないかと期待される。

OSCE における評価の評価者間の差について、村上ら¹³⁾は評価経験が少ない評価者の評価は評価経験をつんだ評価者よりも甘くなる傾向がみられると報告している。しかし、受験者への公平性という観点から、同一内容の試験で結果も同じなら評価は可能な限り一致するように様々なファクターについて検証し適正化を図り、標準化されるべきである。評価者、評価方法、

評価内容、試験時間、それぞれの再現性など多種にわたり検討を加える必要がある。

結 論

OSCE における複数評価者の評価の一致性を検討する目的で本学2003年度歯科医師臨床研修医27名に行った2回のOSCEの結果により一致率とカッパ係数を求めた。

評価する領域別の評価者ペアの評価の一致の度合いは、認知領域、情意領域、精神運動領域の順に高く、情意 認知領域および精神運動領域 認知領域間において有意な差を認めた。

評価者ペア間の評価の一致が著しく低い場合、評価する項目や評価の方法が適切でない場合がある。また、同一の課題においてOSCEの評価者ペアについて観察すると、他のペアと比べ全ての領域で評価の一致が著しく低い評価者ペアが存在し、評価者の適正についても検討する必要がある。

文 献

- 1) Harden RM, Stevenson M, Downie WW and Wilson GM. Assessment of clinical competence using objective structured examination. *Br Med J*. 1975; 22: 447-451.
- 2) Richard KR, 伴 信太郎. カナダの国家試験における Objective Structured Examination (OSCE). *医学教育*. 1998; 29: 9-13.
- 3) Mossey PA, Newton JP and Stirrups DR. Scope of the OSCE in the assessment of clinical skills in dentistry. *Br Dent J*. 2001; 190: 323-335.
- 4) 伴 信太郎, 澤田 司, 田坂佳千, 佐々木宏起, 葛西龍樹, 沸波満, 東 理, 青井一展, 越智則晶, 山本泰博, 伊藤克浩, Kachur EK. OSCE による「臨床入門」実習の評価. *医学教育*. 1994; 25: 327-335.
- 5) 野首孝祠, 小口春久, 古谷野潔, 一戸達也. 臨床実習開始前の共用試験の正式実施を迎えて. *日歯教誌*. 2006; 22: 241-248.
- 6) 藤原 周, 岩堀正俊, 横山貴紀, 岡 敏男, 吉田隆一, 斉藤達哉, 住友伸一郎, 北後光信, 堀田正人, 脇阪孝, 長谷川信乃, 安田順一, 柴田俊一, 安達 潤, 倉知正和. 臨床研修医における客観臨床能力試験について 研修前期におけるOSCEの成績評価法. *日歯教誌*. 2004; 20: 75-84.
- 7) 岩堀正俊, 藤原 周, 横山貴紀, 岡 敏男, 澁谷俊昭, 吉田隆一, 住友伸一郎, 柴田俊一, 北後光信, 松岡正登, 堀田正人, 小川雅之, 森 康志, 倉知正和, 都尾元宣. 卒後臨床研修医の研修に対する意識調査と研修前後のOSCE成績の変化. *日歯教誌*. 2006; 33: 136-143.
- 8) 田口則宏, 小川哲次, 森下真行, 富士谷盛興, 吉野宏, 谷 亮治, 伊藤良明, 田地 豪, 玉本光弘, 渡邊峰朗, 海原康孝, 田口 明, 寶田 貫. 総合歯科医療研修におけるOSCEの標準化 研修前期, 中期, 後期における到達度の推移. *日歯教誌*. 2003; 18: 239-247.
- 9) 相澤文恵, 米満正美, 水城春実. OSCE 評価の信頼性・妥当性と評価項目文に用いる「ことば」の関連性. *日歯教誌*. 2005; 21: 123-132.
- 10) 伊藤孝訓, 青木伸一郎, 大峰浩隆, 河相安彦, 葛西一貴, 金澤英作, 大竹繁雄. OSCE での医療面接における評価の差について ヒューマンエラー分析による検討. *日歯教誌*. 2005; 21: 21-30.
- 11) 宮脇正一, 出口 徹, 村上 薫, 本城 正, 福永智広, 上岡 寛, 吉田登志子, 山本照子. 説明系OSCE開始直後の評価の信頼性. *日歯教誌*. 2007; 23: 299-304.
- 12) 相澤文恵, 岸 光男, 熊谷敦史, 石川義人, 藤沢政紀, 清野幸男, 戸塚盛雄, 米満正美. OSCE における評価の妥当性に関する検討 第1報 複数評価者間の評価の一致度についての分析. *日歯教誌*. 2003; 19: 109-118.
- 13) 村上純子, 竹中秀夫, 堀越 昶, 沢田海彦, 佐藤 貢, 大井洋之, 村上正人, 矢内 充, 竹内 仁, 熊坂一成, 矢崎誠治. 客観的臨床能力試験(OSCE)における医療面接評価の問題点 評価の客観性をよりたかめるために. *医学教育*. 2001; 32: 231-237.