

## 【研究報告】

## 学部4年生対象の成人看護学実習前に実施するOSCEの評価者間の評価得点の一致度

鈴木香苗\*, 百田武司\*, 小川沙苗\*, 岡田淳子\*,  
松本由恵\*, 横山奈未\*, 中信利恵子\*, 植田喜久子\*

## 【要旨】

目的：実習前に実施するOSCEの評価項目・基準を検討するため、OSCE評価者間の評価得点の一致度を明らかにした。

方法：OSCEは、学部4年生対象の成人看護学実習の前に実施し、1名の学生に対して2名の評価者が15項目について評価した。分析は、評価項目毎にカッパ係数を算出し、カッパ係数は、Feinstein (1985)の基準で解釈した。

結果：対象は、評価者ペア13組、学生70名分の評価得点とした。評価者の評価が、『ほぼ完璧な一致』『堅固な一致』は、「予定量、流量の設定」「輸液ポンプの使用前点検」などで、『まあまあな一致』『わずかな一致』は、「輸液の説明」「輸液と患者氏名の確認」「対象者への配慮」であった。

結論：評価の一致度が高い項目は、評価者が学生の行動から明確に判断できるものであり、一致度が低い項目は、評価者の看護観が評価に影響するためにばらつきが生じることが考えられる。「配慮」は、『わずかな一致』であり、評価のあり方を検討する必要性が示唆された。

【キーワード】 OSCE 評価の一致度 カッパ係数

### I. はじめに

A大学では、学部の1年から4年にわたり、実習前及び卒業前にOSCE (Objective Structured Clinical Examination) を行ってきた。OSCEが学生の学習への動機づけになるよう、OSCE評価者(以下、評価者)は、学生が実施した看護について評価表に基づく評価と、フィードバックを行っている。評価は、学生が多様な視点からフィードバックを受けられるように、原則、看護基礎教育を担う教員と臨床看護師である臨床実習指導者(以下、実習指導者)の二者が行っている。

OSCEは、試験であり評価の客観性が重要で、複数の評価者が評価を行う場合、評価の差異が生じることを最小限にする必要がある。歯学分野で、実施されているOSCEにおいて、信頼性、妥当性の担保に関して、各評価項目の点数は評価者の主観に影響を受けることが多いために、評価者間の評価のばらつきが生じることが報告されている(岩堀他, 2009)。本OSCEは、立場の異なる2名の評価者が評価を行うことから、評価者間の評価のばらつきによる客観性の低下が懸念された。

看護教育においては、OSCEの評価の客観性について、教員が評価した評価得点の平均値から得点の低い評価項目の内容を検討している報告がある(内田他, 2008; 浅川, 2006; 鈴木, 2008; 菅野他, 2003)。その他は、評価者間の評価が一致した確率(以下、一致率)を算出し、評価項目について検討したもののみであった(近藤他, 2011)。評価の一致率を単純に見るだけでは偶然に一致した割合(偶然の一致率)の影響が排除できないとされており、評価の信頼性を高めるためには、一致率から偶然の一致率を排除した「一致度(reproducibility)」を指標として用いることが望まれる(Landis and Koch, 1977)。二者の評価の一致度を評価する指標として、実際の一致率が偶然の一致に等しい場合、すなわち、一致が偶然によるものならば0、完全に一致するならば1を示すカッパ係数(kappa statistic, 以下、 $\kappa$ 係数)がある(遠藤, 1998; Kundel and Polansky, 2003)。看護教育におけるOSCEで、複数評価者の評価の偶然を加味した一致度についての報告では、我々が行った研究(百田他, 2013)が、学部2年生の成人看護学実習前に実施するOSCE

\*日本赤十字広島看護大学

の教員と実習指導者の二評価者間の違いによる評価の一致度を検討している。学部2年生対象のOSCEは、学生がOSCEのための技術練習ができるよう、学生に、練習用の課題と評価表を提示している。二評価者間の一致度を算出し、評価基準が具体的なものは一致度が高く、評価者の主観に左右されやすい表現を含む評価基準は、一致度が低い結果が得られている。成人看護学では、学部2年生と学部3～4年生の2度の成人看護学実習を行っており、それぞれ実習前OSCEを実施している。本研究は、我々が行った研究（百田他，2013）とは、異なる学部4年生の成人看護学実習前に実施するOSCEを対象とした。さらに、この学部4年生対象のOSCEは、学生に評価表は提示せず、実際に使用するOSCE課題を3課題提示し、そのうちの1課題のOSCEを受験するという点や、実習グループ毎に、課題について、教員が整備した学習プログラムを活用して学習、技術練習を行っているという点で、我々が行った研究（百田他，2013）とは異なっている。成人看護学では、対象学生や、課題の提示方法の異なる2種類の成人看護学実習を行っており、いずれのOSCEについても評価項目、評価基準を検討する必要がある。そのため、本研究の目的を学部4年生対象の成人看護学実習前に実施するOSCEの二評価者間の評価得点の一致度を $\kappa$ 係数を用いて、明らかにすることとした。

## II. 方法

### 1. 対象

A大学の成人看護学実習の実習前OSCEを受けた学部4年生76名の評価表を分析対象とした。

### 2. 成人看護学実習の実習前OSCEの概要

成人看護学実習において、検査や治療を行う成人を受け持つことが多いため、キュアとケアの統合をめざす看護実践力を育むことをめざした課題とし、実習初日に大学内で実施した。

#### 1) 成人看護学実習の実習前OSCEの課題

目標は、医療機器を使用した患者への看護実践を学習することとし、課題は「輸液ポンプを使用した患者の看護」とした。行動目標は、学生が患者に説明を行い、注射処方箋の内容を理解し、注射薬を確認し、輸液セットを輸液ポンプに適切にセットしたうえで輸液ポンプを操作できること、患者の苦痛や苦悩を理解し、安全かつ安楽であるように何が必要かを考えて輸液療法を実践できるようになることとした。

#### 2) 成人看護学実習の実習前OSCEの評価方法

OSCEの評価表は、容易に客観的評価が出来るよう一文一義とし、課題の実施に必要な技術要素を15の評価項目、45の評価基準から構成した。評価項目の得点は0～3点で、合計30点とした（表1）。

OSCE実施前には、全ての評価者にOSCEの概要、評価者の役割やフィードバックの目的・実施内容・方法、OSCEの評価方法を説明したOSCE評価者ガイドを配布した。そして、OSCEの当日、OSCE課題を説明したうえで、教員が学生役をしたデモンストレーションを実施し、評価項目、評価基準について不明瞭な面を確認する時間を設けた。また、本研究の主旨を説明し、評価過程では、評価者間で一切の相談をしないことを確認した。

#### 3) 成人看護学実習の実習前OSCEの実施方法

OSCEステーションは、より臨床に近い状況を再現した。看護教育用人形を患者とし、閉鎖式輸液セットを用いて静脈留置カテーテルを静脈注射シミュレーターモデル（腕モデル）に挿入し、膀胱留置カテーテルを装着した。そして、ナースコール、輸液ポンプを設置した点滴スタンド、必要物品を運搬するワゴンを準備し、実際に輸液ポンプを使用した輸液療法を開始できるように準備した（図1）。また、学生が患者確認を実際に行えるように、ベッドネーム、ネームバンド、注射処方箋、輸液バッグに貼付するラベルを作成し準備した。さらに、注射処方箋を読解し、輸液総量および時間流量を計算することができるよう、輸液ボトルに薬液を混合した事例を作成した。学生一人あたりのOSCEは、課題文と注射処方箋の読解2分、OSCEの実施10分、フィードバック3分の合計15分間とした。

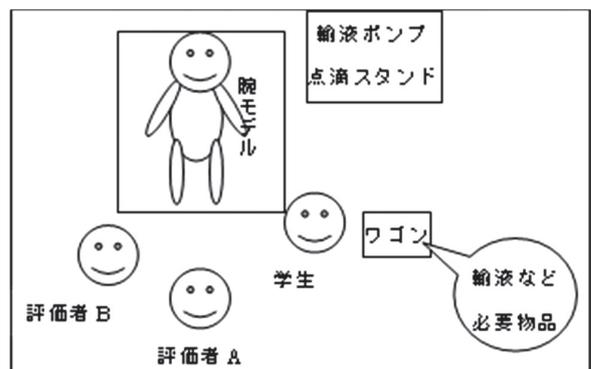


図1 OSCEステーションの配置

### 3. データ収集期間

平成24年5～7月

表1 成人看護学実習Ⅱ 実習前 OSCE 評価表と評価得点の一致度

評価項目	評価基準	配点	評価項目の配点	二評価者間の評価得点の比較			二評価者間の評価得点の一致度		
				教員	実習指導者	P値*	κ係数	一致の評価 <sup>b</sup>	
									平均±SD
1. 輸液ポンプを使用する輸液の説明	輸液ポンプを使用することを説明した	1	1	2	1.33±0.56	1.44±0.67	.276	0.32	まあまあ
	輸液療法の目的を説明した	1	1						
	説明しなかった	0	0						
2. 輸液ポンプの使用前点検	使用前に扉を開け電源を入れて、動作状況を確認した	1	1	1	0.57±0.50	0.67±0.47	.226	0.67	堅固
	確認しなかった	0	0						
3. 注射薬と患者氏名の確認	処方箋と注射薬が合っているか声に出して確認した	1	1	2	1.53±0.61	1.73±0.48	.032*	0.31	まあまあ
	患者のフルネームを声に出しネームバンドで確認した	1	1						
	確認しなかった	0	0						
4. 輸液バックを点滴スタンドにかけ、輸液ポンプの準備	輸液バックを輸液ポンプの真上を避けてスタンドにかけた	1	1	3	2.14±0.87	1.81±0.97	.037*	0.43	中等度
	輸液ルートは輸液ポンプに適正な緩みをもたせて装着した	1	1						
	点滴筒が傾かず、ブローブを滴下ノズルと液面の間に装着した	1	1						
	輸液ポンプに輸液ルートを適切に装着しなかった(点滴ブローブ未装着)	0	0						
5. 予定量、流量の設定	予定量を正しく計算し設定した	1	1	2	1.76±0.62	1.77±0.59	.890	0.83	ほぼ完璧
	流量を正しく計算し設定した	1	1						
	設定できなかった	0	0						
6. 輸液開始前のカテーテル挿入部位の観察	(輸液ポンプ装着前に) 発赤を確認した(口頭で確認)	1	1	3	2.66±0.80	2.66±0.85	1.000	0.52	中等度
	(輸液ポンプ装着前に) 腫脹を確認した(口頭で確認)	1	1						
	(輸液ポンプ装着前に) 疼痛を確認した(口頭で確認)	1	1						
	(輸液ポンプ装着前に) 観察しなかった	0	0						
7. 輸液開始後のカテーテル挿入部位の観察	(輸液開始後に) 発赤を確認した(口頭で確認)	1	1	3	1.09±1.36	1.09±1.37	1.000	0.78	堅固
	(輸液開始後に) 腫脹を確認した(口頭で確認)	1	1						
	(輸液開始後に) 疼痛を確認した(口頭で確認)	1	1						
	(輸液開始後に) 観察しなかった	0	0						
8. 留置カテーテルと輸液セットの接続部の消毒	接続部位を消毒した	1	1	2	1.71±0.59	1.60±0.71	.303	0.46	中等度
	消毒した消毒部位に触れずに接続した	1	1						
	接続時に消毒しない、または接続部位に触れた	0	0						
9. 留置針の固定の確認	穿刺部のテープがはがれていないか確認した	1	1	2	0.39±0.71	0.67±0.93	.042*	0.13	わずか
	留置針とカテーテルの緩みがないか確認した	1	1						
	確認しなかった	0	0						
10. 輸液ポンプの開始	クレンメを開放し輸液ポンプを開始した	1	1	1	0.79±0.41	0.71±0.46	.333	0.66	堅固
	クレンメを開放しなかった	0	0						
11. 輸液ポンプの送液確認	正しく送液されているか確認した	1	1	1	0.66±0.48	0.59±0.50	.387	0.55	中等度
	確認しなかった	0	0						
12. 対象者への配慮	実施中に対象者に配慮して行動をしていた	1	1	1	0.83±0.38	0.96±0.20	.014*	0.07	わずか
	実施中に対象者に配慮をしていなかった	0	0						
13. 対象者の布団を元に戻し、環境の調整	対象者がナースコールを押せる位置に置いた	1	1	3	1.90±1.22	1.76±1.30	.504	0.58	中等度
	対象者にナースコールの位置を説明をした	1	1						
	布団をかけた 布団をめくったまま(ナースコールの位置を整えず) 終えた	1 0	1 0						
14. 看護チームリーダーに実施した援助と対象者の状況を報告	輸液ポンプを設置したことを報告した	1	1	2	0.66±0.72	0.80±0.86	.289	0.62	堅固
	対象者の状態について報告した	1	1						
	報告しなかった	0	0						
15. 輸液の予定量、流量の計算	輸液の予定量を正確に説明した	1	1	2	1.19±0.97	1.03±0.96	.337	0.62	堅固
	輸液の流量を正確に説明した	1	1						
	いずれも正確に説明ができなかった	0	0						
全評価項目の合計点				30	19.21±0.66	19.29±0.60	.889		

a: t検定

b: Feinstein(1985)によるκ係数の解釈

\* P<.05

#### 4. 分析方法

教員と臨床実習指導者の評価者間における評価結果の差異についての分析は、15の評価項目それぞれの平均点を算出し、t検定を行った。次に、1人の学生につき評価を行った2名の評価者の評価の一致度について、偶然性を加味した検討を行うために、15の評価項目毎にκ係数を算出した。データは平均値±標準偏差で表し、有意水準は5%とした。κ係数の値は、 $-1 \leq \kappa \leq 1$  となり、数値が1に近いほど評価者の評価は一致していることを表すが、その解釈は、Feinstein (1985) の基準【0.01~0.20: わずかな (slight), 0.21~0.40: まあまあの (fair), 0.41~0.60: 中等度の (moderate), 0.61~0.80: 堅固な (substantial), 0.81~1.00: ほぼ完璧 (almost

perfect)】とした。解析には、SPSS Ver.20を使用した。

#### 5. 倫理的配慮

OSCE実施前に、OSCEを受験する学生と、本研究の共同研究者である教員以外のOSCEの評価者それぞれに、研究概要、目的、方法、安全性、プライバシーの保護、協力及び辞退の任意性、成績評価などとは一切関係ないことなどについて、文書と口頭で説明し、それぞれに評価表を分析対象とするものの同意を同意書への署名にて得た。なお、研究実施前に日本赤十字広島看護大学研究倫理委員会の承認(No.1116)を得た上で実施した。

### Ⅲ. 結 果

#### 1. 分析対象の概要

同意の得られた学生76名, 評価者21名(教員8名, 実習指導者13名)が採点した評価表152枚のうち, 評価者が教員と実習指導者であった学生70名分, 140枚の評価表を分析対象とした。

評価者の学部3~4年生対象のOSCE評価の経験回数は, 教員が最多11回, 最少2回で, 平均 $5.50 \pm 3.35$ 回, 実習指導者は, 最多4回, 最少1回で, 平均 $1.46 \pm 0.84$ 回であった。

#### 2. 教員と実習指導者の二評価者間の評価得点の比較(表1)

二評価者間の評価得点は, 4つの評価項目で, 差異がみられた。

教員の評価得点が有意に高い項目は, 「4. 輸液バックを点滴スタンドにかけ, 輸液ポンプの準備」で, 教員 $2.14 \pm 0.87$ 点, 実習指導者 $1.81 \pm 0.97$ 点( $p=.037$ )であった。実習指導者の評価得点が有意に高かった項目は, 「3. 注射薬と患者氏名の確認」教員 $1.53 \pm 0.61$ 点, 実習指導者 $1.73 \pm 0.48$ 点, ( $p=.032$ ), 「9. 留置針の固定と確認」教員 $0.39 \pm 0.71$ 点, 実習指導者 $0.67 \pm 0.93$ 点( $p=.042$ ), 「12. 対象者への配慮」教員 $0.83 \pm 0.38$ 点, 実習指導者 $0.96 \pm 0.20$ 点( $p=.014$ )であった。

#### 3. 教員と実習指導者の二評価者間の評価得点の一致度(表1)

二評価者間の評価得点で, 『ほぼ完璧な一致』を示したのは, 「5. 予定量, 流量の設定」0.83であった。『堅固な一致』を示したのは, 「7. 輸液開始後のカテーテル挿入部位の観察」0.78, 「2. 輸液ポンプの使用前点検」0.67, 「10. 輸液ポンプの開始」0.66, 「14. 看護チームリーダーに実施した援助と対象者の状況を報告」0.62, 「15. 輸液の予定量, 流量の計算」0.66であった。『中等度の一致』を示したのは, 「13. 対象者の布団を元に戻し, 環境の調整」0.58, 「11. 輸液ポンプの送液確認」0.55, 「6. 輸液開始前のカテーテル挿入部の観察」0.52, 「8. 留置カテーテルと輸液セットの接続部の消毒」0.46, 「4. 輸液バックを点滴スタンドにかけ, 輸液ポンプの準備」0.43であった。『まあまあ的一致』を示したのは, 「1. 輸液ポンプを使用する輸液の説明」0.32, 「3. 注射薬と患者氏名の確認」0.31であった。『わずかな一致』を示したのは, 「9. 留置針の固定の確認」0.13, 「12. 対象者への配慮」0.07であった。

### Ⅳ. 考 察

#### 1. 教員と実習指導者の二評価者間の評価得点の比較

二評価者間の評価得点は, 教員の方が高い項目は1つで, 実習指導者が高いのは, 3項目であった。その3項目は, 評価基準が一文一義ではなく, 曖昧な表現であった。しかし, 教員の評価得点が高かった項目は, 教員の評価得点が高くなる明確な根拠は明らかではなく, 今後, 検証をしていく必要があると思われる。また, 実習指導者の評価得点が高くなった項目について, 評価得点は, 評価者ができたと判断して, 加点することから, 評価経験が少ない評価者の評価は評価経験を積んだ評価者よりも甘くなる傾向がある(村上, 2010)という報告がある。本研究では, 実習指導者の評価経験の回数は, 教員の評価経験の回数より少なかったことから, 実習指導者の加点の判断が甘くなった可能性があると考えられる。

#### 2. 教員と実習指導者の二評価者間の一致度

『ほぼ完璧な一致』『堅固な一致』を示した評価項目は, 輸液ポンプの操作, 留置カテーテルの挿入部位の観察, 看護チームリーダーへの報告であった。輸液ポンプの操作は, ボタンを押して作動すれば, 予定量, 流量が数字で表示するため, 実施の有無の判断ができる。また, カテーテル挿入部位の観察は, 学生に発赤, 腫脹, 疼痛の確認を口頭で行うことを求めており, 言語で表現しなければ加点しない。さらに, 評価基準を3つの内容の確認に対して, 設定していることから, 学生がどのような行動によって確認をするのかが具体的であった。我々が学部2年生を対象に行った研究(百田他, 2013)も, 『堅固な一致』を示した項目は他の項目よりも具体的な基準が示されていたとの結果が得られており, 評価基準は, 行動レベルまで具体的にすることで評価者の判断が一致したと考えられる。そして, 看護チームリーダーへの報告については, 学生が行う報告の内容を簡潔な文章にしたことが一致度を挙げたのではないかと考えられる。

『まあまあ的一致』『わずかな一致』を示した評価項目は, 「説明する」「確認する」という行動の幅が広いものであり, 学生の言動を見て, できたと判断する基準が曖昧であった。複数評価者の評価の一致度は, 評価項目において, 曖昧な表現があるものは, 低くなる傾向があることが明らかになっている(森戸, 原, 梅田, 野元, 高崎, 2009; 菅沼他, 2005; 相澤, 米満, 水城, 2005)。評価項目を作成する際は, 評価者の一致度を上げるため, 行動レベ

ルの明瞭な表現を用いた評価基準が必要であることが示唆された。

一致度が最も低かった評価項目は、対象者への配慮であった。認知領域や情意領域の項目は、評価者の経験や価値観などが評価に影響を与える可能性が高い(中村, 2011, p.87)ことが示唆されており、配慮する行動は、評価者が大切にしている看護観が評価に影響するため、ばらつきが生じやすいことが推測される。また、配慮といった情意領域は一側面のみで評価できるものではない事、さらには看護職に求められている役割期待が無限定的にあることなどから、具体的な評価項目の設定が難しいことが明らかになっており(灘, 2004)、配慮については、評価項目を特定すること、さらに点数化することには限界があると考えられる。

よって、配慮する行動は、OSCE 評価得点で評価するのではなく、実施後の評価者からのフィードバックで学生の気付きを促したり、学生同士で看護師役・患者役を経験するなど、配慮とは何かを多面的に考えるようにしたりと教員が工夫して企画していく必要性が示唆された。

## V. 結 語

1. 教員と実習指導者の二評価者間の評価得点は、15の評価項目の評価得点のうち、実習指導者の評価得点が高いものが3つであった。OSCE の評価経験回数が少ない実習指導者は、加点の判断が甘くなり、評価得点が高くなった可能性がある。
2. 教員と実習指導者の二評価者間の評価得点の一致度が『ほぼ完璧な一致』、『堅固な一致』を示したのは、評価基準が具体的に行動レベルとなっており、客観的なものであった。『まあまあ的一致』、『わずかな一致』を示した評価項目は、評価基準が曖昧な表現であった。
3. 配慮する行動は、OSCE 評価得点で評価するのではなく、実施後の評価者からのフィードバックで学生の気付きを促したり、学生同士で看護師役・患者役を経験するなど、配慮とは何かを多面的に考えるようにしたりと教員が工夫して企画していく必要性が示唆された。

## 謝 辞

本研究を行うにあたり、ご協力いただきました学生、実習指導者の皆様に心より厚くお礼申し上げます。

本研究は、平成24~25年度日本赤十字広島看護大学共同研究費(研究代表者鈴木香苗)の助成を受け

て実施したものであり、結果の一部を日本看護学教育学会第23回学術集会において発表した。

## 文 献

- 相澤文恵, 米満正美, 水城春実 (2005). 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. 日本歯科医学教育学会雑誌, 21(2), 123-132.
- 浅川和美 (2006). 学生の技術力を身につけるための臨地実習教育全領域での OSCE (客観的臨床能力試験) による技術習得度の評価, 看護展望, 31(2), 75-81.
- 遠藤和男 (1998). 医学における統計学の応用について(第6編)診断における統計学の応用について. 新潟医学会雑誌, 112(3), 135-143.
- Feinstein, RA (1985). Clinical epidemiology. The Architecture of clinical research, WA Saunders, Philadelphia, 184-186.
- 百田武司, 鈴木香苗, 小川沙苗, 岡田淳子, 松本由恵, 中信利恵子, 池田奈未, 植田喜久子 (2013). 学部2年次対象の成人看護学実習前に実施する OSCE の, 評価者間の違いによる評価の一致度に関する検討. 日本赤十字広島看護大学紀要, 13, 1-8.
- 岩堀正俊, 小川雅之, 広瀬俊, 米田博紀, 大森俊和, 都尾元宣, 藤原周, 住友伸一郎, 村松泰徳 (2009). OSCE の評価者の違いにより評価の一致性に関する検討. 岐阜歯科学会雑誌, 35(3), 160-166.
- 近藤智恵, 市村久美子, 伊藤香世子, 高橋由紀, 沼口知恵子, 黒田暢子 (2011). OSCE における教員間の評価の差異と課題. 茨木県立医療大学紀要, 16, 1-11.
- Kundel HL, Polansky M (2003). Measurement of observer agreement. Radiology, 228(2), 303-308.
- Landis JR, Koch GG (1997). The measurement of observer agreement for categorical data. Biometrics, 33, 159-174.
- 森戸麻美, 原早苗, 梅田伸威, 野元和夫, 高崎正彦 (2009). OSCE 評価の一致性と模擬患者における評価に関する検討. 全日本鍼灸学会雑誌, 59(1), 22-29.
- 村上純子, 矢内充, 竹中秀夫, 佐藤貢, 沢田海彦, 村上正人, 竹内仁, 大井洋之, 矢崎誠治 (2001). 客観的臨床能力試験(OSCE)における医療面接評価の問題点 評価の客観性を高めるために. 医学教育, 32(4).231-237.
- 灘久代 (2004). 看護学生の臨地実習における行動

- 変化と態度育成. 島根県立看護短期大学紀要, 9, 33-38.
- 中村恵子 (2011). OSCE の評価方法看護における知識・技術・態度をどのように評価するか評価方法のプロセス. 看護 OSCE(第1版). (p87). 東京. メジカルフレンド社.
- 菅野こずえ, 大久保祐子, 里光やよい, 豊田省子, 亀田真美, 野中静, 田口ヨウ子 (2003). チーム医療・看護とコメディカルの教育72 看護学 OSCE における評価者間の評価の信頼性の検討 - 信頼性に影響する要因 -. 医学教育, 34, 43.
- 菅沼岳史, 角田左武郎, 長谷川篤司, 木村裕一, 宮下元, 天野均, 塚崎弘明, 五島衣子, 岡野友宏, 佐野司 (2005). 客観的臨床能力試験 (OSCE) における評価の一致度について. 日本歯科医学教育学会雑誌, 21(1), 3-9.
- 鈴木玲子 (2008). 看護技術の教育・評価成人看護領域における OSCE の展開—看護実践力の向上につながる評価のあり方. 看護展望, 33 (3), 283-288.
- 鶴木恭子, 宮崎みち子, 内田雅子, 大野夏代, 清水光子, 田中広美, 藤井瑞恵, 三上智子, 中村恵子 (2012). OSCE 実施時に評価者が「採点しにくい」と感じた理由—1・2年生 OSCE 評価者アンケートの結果から—. 札幌市立大学研究論文集, 6(1), 11-18.
- 内田倫子, 土屋八千代, 赤星成子, 山田美由紀, 緒方昭子, 奥祥子 (2008). 成人看護学における OSCE の試み. 南九州看護研究誌, 6(1), 55-61.

# Degree of Inter-evaluator Agreement for an OSCE Held Prior to an Adult Nursing Practicum Aimed at College Seniors

Kanae SUZUKI\*, Takeshi HYAKUTA\*, Sanae OGAWA\*, Junko OKADA\*,  
Yoshie MATSUMOTO\*, Nami YOKOYAMA\*,  
Rieko NAKANOBU\*, Kikuko UEDA\*

## Abstract:

**Objective:** In order to examine the evaluation items and standards of an OSCE held prior to a practicum, this study clarified the degree of inter-evaluator agreement of evaluation scores.

**Methods:** The OSCE was held prior to an adult nursing practicum for college seniors. Two evaluators per student evaluated 15 items. For analysis, the kappa coefficient was calculated for each item. Interpretation of the kappa coefficient was based on Feinstein's (1985) standards.

**Results:** The focus of analysis was the evaluation scores given by 13 pairs of evaluators to 70 students. Almost perfect agreement or solid agreement was achieved for "setting the planned dosage or flow volume" and "checking before use of the transfusion pump." Moderate and minor agreement were achieved for "explanation of transfusion," "confirmation of transfusion with patient's name," and "consideration for subjects."

**Conclusion:** Items with a high degree of agreement between evaluators were those for which the evaluators could make clear judgments based on student actions. For items with a low degree of agreement, it appears that the evaluators' own nursing philosophy influenced the evaluation, leading to variation. The minor agreement for "consideration" suggests the need to examine the evaluation methods further.

## Keywords:

agreement in OSCE evaluations, kappa coefficient

---

\* Japanese Red Cross Hiroshima College of Nursing

