

薬学 OSCE での患者対応課題における 評価方法に関する検討

窪田愛恵^{1,4}, 矢野義孝^{*1,4}, 森本 剛², 高田香織², 蔵本伸生²,
関 進², 前田祐子³, 赤池昭紀¹, 平出 敦²

京都大学薬学研究科¹, 京都大学医学研究科医学教育推進センター²,
京都大学大学院医学研究科人間健康科学系専攻³,
京都大学薬学研究科統合薬学フロンティア教育センター⁴

Analysis of Evaluation System for Communication Skills in Pharmaceutical OSCE

Yoshie Kubota^{1,4}, Yoshitaka Yano^{*1,4}, Takeshi Morimoto², Kaori Takada²,
Nobuo Kuramoto², Susumu Seki², Yuko Maeda³, Akinori Akaike¹ and Atsushi Hiraide²

Graduate School of Pharmaceutical Sciences, Kyoto University¹,
Center for Medical Education, Kyoto University²,
School of Health Sciences, Faculty of Medicine, Kyoto University³,
Center for Integrative Education of Pharmacy Frontier,
Graduate School of Pharmaceutical Sciences, Kyoto University⁴

〔 Received May 23, 2008 〕
〔 Accepted August 4, 2008 〕

For the objective structured clinical examination (OSCE) used in pharmaceutical education, it is important to establish valid evaluation systems, especially for the evaluation of student communication skills in medical interviews. In the present study, we examined the evaluation scores obtained in the 2nd OSCE trial at Kyoto University, and discuss the OSCE evaluation system for communication skills. The results of kappa analysis, which was used to determine the proportion of agreement between the scores of two instructors, suggested that there was a large variation in the proportion of agreement, the extent of which depended on the type of evaluation questionnaire and the amount of ambiguity in it. In general, there was a higher proportion of agreement among the items of closed questions, and among open question items for evaluation of student behavior there was a lower degree of agreement. These results were confirmed by a review of recorded images and voices. Based on our results and considering that medical interview skill is evaluated through a conversation between a student and a standardized patient, we conclude that it is not practical to determine the evaluation standard for each questionnaire too precisely. We also feel that summarized scores would be preferable for the evaluation of student communication skill in OSCE.

Key words — OSCE, communication, medical interview, kappa analysis, proportion of agreement, standardization

緒 言

薬学部における薬剤師養成のための6年制教育が開始されたが、その中で2年後に正式実施となる薬学部共用試験のうち、特に技能・態度を評価する試験である「客観的臨床能力試験(Objective Structured Clinical Examina-

tion: OSCE)」（日本薬学会薬学教育改革大学人会議 OSCE 内容・体制委員会, http://pc5.phcat-unet.ocn.ne.jp/about_osces.html）の正式実施に向けたトライアルが全国の大学で実施されている¹⁻³⁾。OSCEで実施される課題は5領域から6課題と予定されているが⁴⁾、その中でもコミュニケーションに関する技能・態度を問う医療面接課題、すなわち患者・来局者対応および情報提供の課題におい

* 京都市左京区吉田下阿達町 46-29; 46-29, Yoshida-Shimoadachi, Sakyo-ku, Kyoto-shi, Kyoto, 606-8501 Japan

ては、評価者や模擬患者の養成を含む評価体制の整備が重要課題の一つとなっている。

6年制教育では長期の実務実習が義務化されており、実習先となる医療現場での医療関係者および患者とのコミュニケーションは学生にとって非常に重要となる。このため、医療面接に関しても他の技能実習と同様に事前実習が重要であり⁵⁾、その成果がOSCEで評価されることになるが、その課題内容は模擬患者を相手に行う会話形式のロールプレイを通じて受験生を評価するものであり、調剤などの実技課題と比べて模擬患者が存在し人と人とのコミュニケーションのなかで学生の技能・態度を評価することから、実技課題とは異なる評価手法が必要と考えられる。また、受験生と模擬患者との一連の会話のなかで適切なコミュニケーションが行われ、かつ受験生が必要事項を伝達あるいは聞き取りできたかどうかを判断する必要があるため、評価者が評価項目を順番に確認するだけでは十分にコミュニケーションに関する本質的な評価とはならないことが懸念される。医学教育においてもOSCEにおける評価体制や評価方法の信頼性に関する検討が実施され⁶⁾、また、模擬患者による評価の重要性についても認識されている⁷⁾。しかしながら、現在までの薬学OSCEトライアルにおいては体制づくりや人員の確保といった点に目的が集中しており、OSCEトライアルの結果を精査し評価システムや評価方法について検証した報告は少ない。

そこで、京都大学薬学部で実施した第2回OSCEトライアルにおいて医療面接課題で得られた情報を整理し、評価方法の問題点とその原因を探ることを目的とした検討を実施した。

方 法

京都大学薬学部において2008年3月1日に実施したOSCEトライアルでの結果を用いた。このトライアルでは京都大学薬学部学生(平成17年度入学の3年生)29名と大学院生2名の合計31名を受験生として、6ステーション(初回面接(薬局での患者応対)、散剤調剤、水剤調剤、調剤鑑査、注射薬混合、薬局での薬剤交付(情報提供)の6課題)からなる試験を実施した。

患者応対に関する技能・態度を問う「初回面接」と題したステーションでは、評価者として医学部教員5名と薬学部教員5名を配置し、各学生に対し各群より1名ずつが評価を担当した。医学部教員群は医療コミュニケーション教育を担当している教員4名および病院で実務教育を行っている薬剤師1名からなる。また、模擬患者として薬学部教員3名、および医学教育において模擬患者のトレーニングを積んでいる一般市民(京都大学模擬患者の会)5名が担当した。

課題の内容は本学で独自に設定したシナリオに従って実施した。具体的には「自転車で転んで右手首を捻挫しシップと痛み止め錠剤を処方された患者が初めての薬局に来局した場面」とした。なお、今回のトライアルでは、各質問項目に対する評価結果をもれなく得るために「症状」、「経過」、「既往症」といった聞き取りすべき項目一覧を受験生に提示し、受験生はそれを参照し必要であれば記録を取りながら面接を進めるといった形をとった。5分の面接実施時間で、課題進行中はすべての受験生の試験風景を画像・音声ともに記録した。学生に対する評価は、個別の評価項目(評価者20項目、模擬患者13項目)と概略評価とによって実施した。

個別の評価項目については、原則として薬学共用試験センターの方針に従った。評価者・模擬患者それぞれの評価表における評価項目の概要を表1に示した。各評価項目について、評価者が「できた」と判断した場合にスコアとして1を、「できなかった」と判断した場合にはスコアとして0を選択するチェックリストを利用した。概略評価は、現時点の薬学OSCEトライアルにおいて一般的に用いられている6段階評価を、評価者・模擬患者ともに用いた。

個別の評価項目に関する解析は、各評価項目で「でき

表1. 「初回面接」課題における個別の評価項目
評価者用

導入段階
1: 適切な挨拶, 2: 自己紹介, 3: 患者氏名確認
4: インタビューの説明と同意取得
情報の収集(確認項目)
5: 症状の部位と程度, 6: 症状の経過, 7: 既往症,
8: 他科受診の有無, 9: 使用中の薬, 健康食品,
10: 喫煙, 飲酒, 11: アレルギー歴(食品など),
12: 副作用歴
締めくくり
13: しめくくりの言葉
コミュニケーション全般
14: 身だしなみ, 15: 姿勢, ふるまい, 16: アイコン
タクト・顔の向き, 17: 声の大きさ, スピード, 音調,
18: 言葉遣い, 19: 傾聴, 20: 共感の言葉, 態度

模擬患者用

あいさつ
1: 最初の挨拶, 2: 目線, 3: しめくくりの言葉
聞き取り技能
4: 判りやすい説明, 5: 患者の言葉を正しく理解, 6:
患者の質問への対応
態度, コミュニケーション全般
7: 身だしなみ, 8: 姿勢, ふるまい, 9: アイコン
タクト・顔の向き, 10: 聞き取りやすい, 11: 言葉遣
い, 12: 傾聴, 13: 共感の言葉, 態度

た(1)」と評価された項目数の全項目数に対する割合を算出し「個別的評価項目得点率」と定義しその度数分布を得た。概略評価に関してはスコアの度数分布を算出し、評価者、模擬患者それぞれについてグラフ化した。さらに、模擬患者の概略評価スコアと各受験生につき2名の評価者による概略評価スコアの平均値との関係について、模擬患者の属性(教員あるいは一般市民模擬患者)で調整した線形重回帰モデルを用いた解析を実施し、概略評価への影響因子を検討した。次に、個別的評価項目について各受験生に対する各評価項目の一致度を検証するため κ 解析⁹⁾を行った。評価者は薬学部教員とコミュニケーション教育を担当している医学部教員とに分け、 κ 係数によって2群の一致度を調べた。 κ 係数による一致度の判定は、 $\kappa < 0.0$:一致なし、 $0.01 < \kappa < 0.20$:わずかに一致、 $0.21 < \kappa < 0.40$:やや一致、 $0.41 < \kappa < 0.60$:だいたい一致、 $0.61 < \kappa < 0.80$:相当な一致、 $0.81 < \kappa < 1.0$:ほぼ完全に一致、とした⁹⁾。なお、評価項目が空欄となっている場合には解析から除外した。

また、すべての受験生について、初回面接課題の録画面像および音声をもとにその会話の内容と評価結果とを比較検討した。

線形モデルによる解析、および κ 解析は統計解析ソフトウェア jmp(SAS Institute Japan(株))により行った。

結 果

評価の全体像を把握するために、図1に評価者10名によるのべ62回の個別的評価項目の得点率および概略評価スコアの度数分布を示す。また、模擬患者8名による評価の分布を比較検討のために示した。個別的評価項目得点率の分布に関しては、評価者では80~90%の得点率が最も多かった。これに対し模擬患者による評価では90%以上の得点率が最も多かった。また、模擬患者の評価では分布が広がっており、6段階で評価する概略評価においても同様の傾向がみられた。概略評価については評価者が6段階で4とつけた例が最も多く、模擬患者では5とつけた例が最も多かった。一方で、模擬患者のほうが評価者に比べて合格境界付近(概略評価2以下を不合格と設定した)の評価が厳しくなる傾向がみられ、不合格と判定した模擬患者はのべ5名であった(図1)。評価者が概略評価で2以下とした例は1名であったが、同じ受験生を評価したもう一人の評価者の概略評価は4であり、異なった評価結果が得られていた。以上のことから、評価者と比較して模擬患者のほうが学生に良い評価を与える割合が高いものの、一方で不合格と判定する場合もあり、模擬患者のほうが評価結果に広がりが見られた。

図2に概略評価スコアに対する線形モデル解析の結果

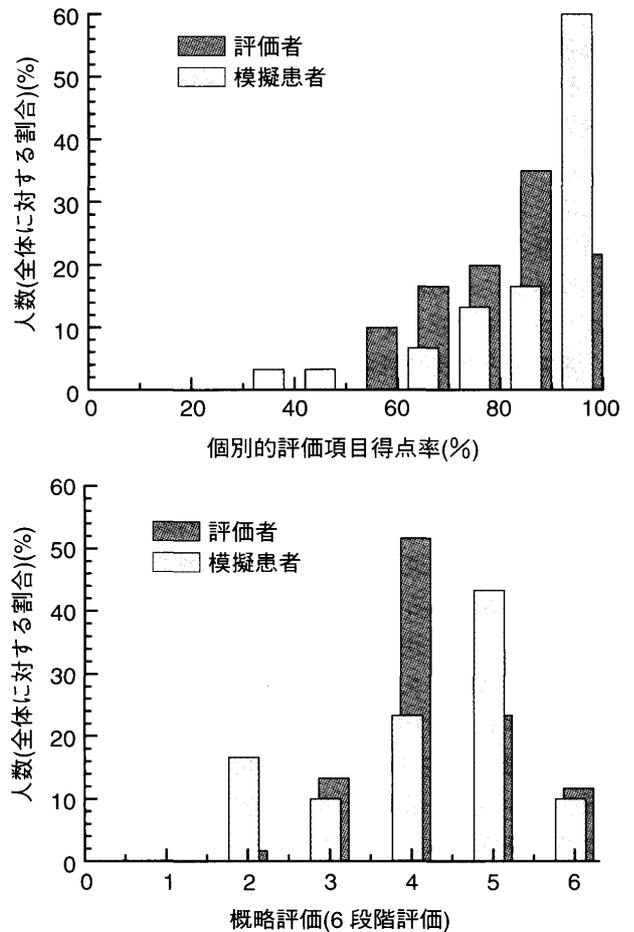


図1. 各個別的評価項目で「できた」と評価された項目数の総項目数に対する割合(個別的評価項目得点率)(上図)と概略評価(6段階評価)スコアの度数分布(下図)

を示す。一人の学生に対する評価者の概略評価(2名の評価者の平均値)と模擬患者の概略評価とは回帰係数が $0.92(p=0.0002)$ とほぼ一致していた。模擬患者の属性による影響を調べた結果、薬学部教員による模擬患者は一般市民模擬患者よりも平均的に 0.69 ポイント低く($p=0.06$)評価する傾向があった。さらに個々の評価結果をみると、図1で述べた概略評価で不合格と判定される2と評価した模擬患者のべ5名(実際には1名×1回、2名×2回)は全員が薬学部教員であり、薬学部教員による模擬患者が一般市民模擬患者と比較して低く評価する傾向はこれらのべ5名の評価結果によることが明らかであった。これは図1でみられた模擬患者全体での評価値の分布の広がりとも関係する。

個別的評価項目について一致度に関する κ 解析の結果を表2に整理した。一致度の目安となる κ 係数は -0.15 から 0.83 に分布し、特に κ 係数が負の値をとり一致度が悪かったと判断できる項目は「8:他科受診の有無($\kappa = -0.051$)」, 「9:使用中の薬, 健康食品($\kappa = -0.15$)」,

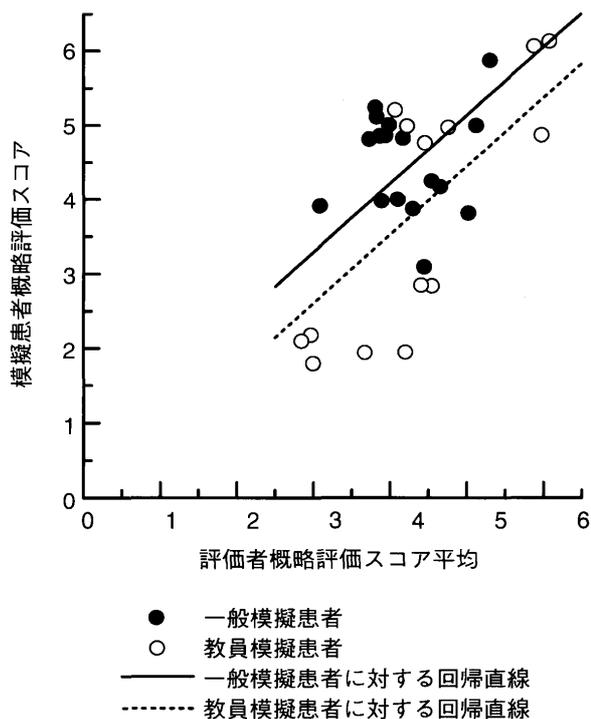


図2. 模擬患者の概略評価(6段階評価)スコアと評価者による概略評価(6段階評価)スコアの平均値との関係
直線は線形重回帰モデル解析の結果をもとに描画した。

表2. 各個別の評価項目の一致度に関する κ 係数

評価項目番号	κ 係数
1(挨拶)	0.37
2(自己紹介)	0.83
3(患者氏名)	0.76
4(説明と同意)	0.42
5(部位と程度)	NA
6(経過)	0.21
7(既往症)	0.44
8(他科)	-0.051
9(薬, 食品)	-0.15
10(喫煙, 飲酒)	NA
11(アレルギー)	NA
12(副作用)	NA
13(しめくり)	0.36
14(身だしなみ)	-0.062
15(姿勢)	0.30
16(目線)	0.61
17(声)	0.56
18(言葉)	NA
19(傾聴)	0.28
20(共感)	0.48

*: NA(Not Applicable; 評価結果がすべて「できた」あるいはすべて「できなかった」となる項目であったため解析不能)

「14:身だしなみ($\kappa = -0.062$)」であった。これらの項目に着目して録画画像と音声とを分析すると「8:他科受診の有無」,「9:使用中の薬, 健康食品」については、例えば他科受診の確認と同時にそこで処方されている薬についても確認し、その結果として健康食品についての確認を忘れるといったように、一連の会話の流れのなかで他の項目とひとまとめにして触れている場合が多くみられ、そのような場合に評価者間で評価結果がわかれたものとなっていた。また、「14:身だしなみ」については評価マニュアルを検証したところ「清潔な白衣, 名札, 髪, 爪, 貴金属を身につけない, 服や履物に問題がない, などをチェックし, 気になる点があれば無得点とする」と記載しており, 複数の外見的項目をまとめて評価する必要があった。また、「気になる点がある」の判断基準には評価者の主観的印象が求められていた。

一致度が比較的大きかった項目として κ 係数が 0.61 以上のものは「2:自己紹介」,「3:患者氏名の確認」,「16:アイコンタクト・顔の向き」であり, 確認が一義的に実施できる項目では画像による検証からみても適切に評価され一致度も高かった。会話においては自己紹介や患者氏名の確認などの単純なやりとりや, 喫煙歴の確認といった閉鎖型質問では一致度は高い傾向があり, 一方で「5(部位と程度)」,「6(経過)」,「7(既往症)」といった開放型質問では一致度が低い傾向にあった。「11(アレルギー)」,「12(副作用)」に関してはほとんどの受験生で「できた」と評価されているものの患者の発話を十分に解釈していなかった例もあった。表3に受験生と模擬患者との実際の会話の例を示す。この例のように, 患者が薬による副作用について発話したにもかかわらず再び副作用に関する質問を行い, 単にアレルギー, 副作用と項目を順番に機械的に質問した例がみられた。ただし, この例の場合には, 模擬患者がアレルギーに関する質問を受けたにもかかわらず副作用に関する回答までを発言しており, 模擬患者の対応を統一させるという点からは問題もある。公平な評価を行うためには, 模擬患者の養成や対

表3. 「アレルギー」および「副作用」の聞き取りに関するある受験生と模擬患者との会話内容

薬剤師: 「何かアレルギーはありますか?」
患者: 「特にアレルギーはないんですが, 前に一回, 風邪のときかな, 薬をもらってそれでちょっとふらふら, フラついたので, はい. 物をたべて蕁麻疹でとか, そんなことはありません。」
薬剤師: 「何か薬で副作用にあわれた経験は?」
患者: 「今, 言いましたように, なんの薬か名前は忘れましたが, 風邪の薬で。」
薬剤師: 「そうですね, 他に聞きたいこととかあればぜひ。」

応マニュアルの作成においてこのような事例が生じないようにすることも必要と考えられる。

一方、概略評価については、評価者間で評価スコアが完全に一致した例は31名の受験生中19例(61%)であり、 κ 係数は0.45であったが、評価スコアが2以上異なった例は2例のみ(記載がなかった例、および概略評価が「2」および「4」となった例)であり、評価スコアの差が±1の範囲内で「一致」と考えると一致率は29/31=94%、 κ 係数による一致度は0.86であった。

なお、医学部教員群と薬学部教員群との間の「個別の評価項目の合計点」について、Spearmanの順位相関係数は0.762と得られ、 $p < 0.01$ で有意な相関が認められた。また、すべての評価者における「個別の評価項目の合計点」と「概略評価スコア」についてSpearmanの順位相関係数は0.629と得られ、 $p < 0.01$ で有意な相関が認められた。

図3に、ある受験生での課題実行時の会話内容および会話の経過時間、さらに各評価項目に対する2名の評価者の評価結果とそれに関連する会話部分との関連を示した。評価者は数秒間の会話の内容から評価項目1~13それぞれに対する判定を行い、また、評価項目14~20で与えられるコミュニケーション全般の評価には試験全体を通しての観察が必要であり、画像および音声の検証結果から同時に複数因子に関して評価を求められている繁忙な評価の状況が明らかであった。

考 察

線形モデル解析の結果から、6段階の概略評価スコアに関しては評価者と模擬患者との間で平均的にはほぼ同じスコアを与えるものと考えられた。しかし、模擬患者の属性による影響を調べた結果、薬学部教員による模擬患者は低く評価する傾向があることが示された。教員では患者視点に加えて教育者視点から評価する部分があり、それが評価の一部に影響したことも理由の一つとして考えられる。特に概略評価で2(不合格と判断)とした模擬患者のすべてが薬学部教員であったが、この結果に関しては例数が限られているが教員が模擬患者役を担当する場合の一つの問題点を示唆するものとも考えられ、今後注目して検討を続ける必要がある。OSCEにおいて医療面接課題を実施する目的の一つは、実務実習で学生が一般人患者と触れ合うために適切な技能・態度を持ち合わせていることを確認することであり、この点から考えるとOSCEではよくトレーニングされた一般人が模擬患者となるべきである。教員が模擬患者を行うのであれば、その養成にあたっては一般市民模擬患者との交流の機会を増やすなど、一般市民模擬患者としての役割が果たせるような工夫が必要と考えられる。

一方、個別の評価項目については、全体的にみて自己紹介や患者氏名の確認などの単純な会話や、喫煙歴の確認といった閉鎖型質問では一致度は高い傾向がみられたものの、「5(部位と程度)」、「6(経過)」、「7(既往症)」といった開放型質問では一致度が低い傾向にあり、今後の評価のあり方を考える意味で課題となる。一致度が低い原因として受験生が単に「質問したか」を重視するのか、あるいは「正しく聞き取れたか」を検証するのかという視点の相違が評価者間にあったことが考えられる。その点では二値形式のチェックリストでは判断に迷ってしまう。また、挨拶や締めくくりの言葉など、一見単純ではあるが判断基準が明確でない項目での不一致も浮き彫りにされた。評価者は発言したかどうかだけでなく非言語的あるいは準言語的な要因についても評価しようとする傾向があり、その結果として評価結果がわかれたと考えられる。この傾向はOSCEの目的から考えても自然であると考える。同様にこうした要因の関与が大きい「15(姿勢、ふるまい)」~「20(共感の言葉、態度)」に関しては κ 係数からみた一致度は悪くはないもののばらつきがみられた。ただし、「16(目線)」、「17(声)」に関しては評価のポイントが限定されており一致度は比較的良好になったものと考えられた。

次に、用語として混乱を招く可能性のある項目として「11(アレルギー)」、「12(副作用)」について考察した。これらの言葉は「アレルギーの伴う副作用」のように意味が重なりあう部分があり、一般患者にとっては判りにくいところがある。現実的には患者から聞き取った内容をもとに、医療者が「アレルギー歴」あるいは「副作用」と判断するものである。録画画像を確認したかぎりでは受験生はそのまま「アレルギー歴の確認」、「副作用歴の確認」といった項目を質問として発言した例が多く、個別の評価で「できた」と判断されていてもコミュニケーションとしては問題があると考えられた。より適切なコミュニケーション手法としては「今まで薬で何か困ったことはないか?」、「蕎麦や卵、鯖などの食品や虫刺され、花粉などによるアレルギーはないか?」というように患者に対応しながら具体的に質問できるようになること、さらに単に質問するだけでなく、相手の話を聞いて自分で解釈し相手に確認できるようになること、このような事前教育が必要であろう。

以上、考察したように、OSCEにおいて特に面接課題では模擬患者との生の会話を評価するため評価項目によっては評価者間で高い一致度が得られないことがわかった。その理由としては、①リスト化された個別の評価項目の順序どおりに会話が進むとは限らず、また、複数の評価項目に対する受験生の反応が同時に出現することがあり、評価者が見逃したり正解と判断しない可能性がある、②評価項目によっては統一的基準を設けること

時間	会話内容	評価項目	評価者スコア	
0:00	薬剤師：こんにちは。	1 (挨拶)	0	1
0:01	患者：失礼します。			
0:04	薬剤師：「田村じゅん」さんでよろしいでしょうか。			
0:05	患者：はい、そうです。よろしくお願いします。			
0:09	薬剤師：よろしくお願いします。それではおかけ下さい。	2 (自己紹介)	1	1
0:11	患者：ありがとうございます。	3 (患者氏名)	1	1
0:14	薬剤師：薬剤師研修生の〇〇と申します。よろしくお願いします。			
0:17	患者：よろしくお願いします。			
0:21	薬剤師：ではまず、あの、症状の確認からさせていただきます。	4 (説明と同意)	0	0
0:26	患者：はい。	5 (部位と程度)	1	1
0:28	薬剤師：どのあたりが痛むんでしょうか。	6 (経過)	1	1
0:30	患者：あの、利き手の手首って言うんですかね、ここがちよつと腫れてて・			
0:39	薬剤師：大変ですね。どのくらい痛みますか。			
0:43	患者：痛みより、腫れの方が気になりまして・			
0:47	薬剤師：わかりました。お医者さんはどのようにおっしゃってましたか。			
0:54	患者：ええっと、とりあえずは痛み止めと湿布で様子見ましようかということ・			
0:59	薬剤師：わかりました。湿布がでてますね。			
1:01	患者：はい。			
1:03	薬剤師：わかりました。他に別のご病気とかなさってるとかはありますか。	7 (既往症)	1	1
1:08	患者：別な病気といますと。			
1:11	薬剤師：他に・	8 (他科)	0	1
1:12	患者：内科とかそういうことですか。			
1:13	薬剤師：はい、そうです。こちらのお医者さん以外に・			
1:16	患者：ここは整形外科だったんですけど、内科とか他の病院の方で・ ええっと、それはね、去年の職場の健診で血圧が高いと言われたもんで、まあ、近くの内科に行 った時に薬を飲んで方が良いつことで、薬をいただいているんですけど。	9 (薬、食品)	0	1
1:38	薬剤師：あ、そうですか。わかりました。ありがとうございます。 それからこちらでは血圧のお薬をもらっているということですね。			
1:47	患者：はい。			
1:49	薬剤師：それ以外にご自身で健康食品などを使われているということはありますか。			
1:53	患者：健康食品は使ってないですけど。			
1:55	薬剤師：そうですか、わかりました。それでは、生活状況についてお伺いしたいんですけども、 おタバコなどお吸いになられますか。	10 (喫煙、飲酒)	1	1
2:03	患者：たばこはすいませんけど。			
2:05	薬剤師：お酒は飲めますか。			
2:08	患者：お酒、ビール1本くらいかな。			
2:11	薬剤師：一日に？			
2:12	患者：そうですね、夕方位に。			
2:14	薬剤師：わかりました。アレルギーなどはありますか。	11 (アレルギー)	1	1
2:18	患者：アレルギーは、うーん、病院の先生の方からはこれといっは言われてはいないんですけ ど。	12 (副作用)	1	1
2:28	薬剤師：わかりました。今までに何かお薬を飲んで副作用がでたとかそういうことはありますか。			
2:34	患者：これは私の憶測なんですよね。あの、ちよつと風邪をひいたときに内科でもらった薬が、 風邪だったんですね、その時にふらつとしたから、あれはちよつと薬のアレルギーかなって思う くらいで、それは先生にどうこう言ったわけじゃなくて、自分の憶測です。			
2:59	薬剤師：はい。わかりました。もし、またそういうことがありましたら、何か副作用かもしれな いので薬剤師なりお医者様へお伝えください。			
3:08	患者：はい。			
3:10	薬剤師：では以上でインタビューが終わりましたので、今から薬を、こちらの薬を用意するの でもうしばらくお待ち下さい。	13 (しめくり)	1	1
3:24	患者：はい。わかりました。			
3:28	終了			

図3. ある受験生の課題実行時の会話内容と各個別の評価項目との関連

左欄に会話開始からの経過時間(分:秒)を示した。また、各個別の評価項目について2名の評価者評価対象となると考えられる発言箇所との関係を矢印で結んだ。評価結果が一致していない項目についてはその原因と考えられる事実を示した。

に限界があり、それぞれの評価者の判断に任せざるを得ないものがある、ことが示された。

一方で、概略評価においては評価スコアの差が±1の範囲内で「一致」と考えた場合に比較的高い一致度が得られた。この値と個別的评价項目のκ係数と単純に比較できるものではないが、一致度が大きく異なり得る各評

価項目の合計点を重視するよりもこれらの評価結果は合否判定において特に「否」とする場合の根拠として、またフィードバックの材料として活用し、現状では概略評価による判定を重視することが現実的と考える。

今回、課題風景を録画することでその画像および音声からいくつかの考察を行うことができた。さらに「2(自

己紹介)や「3(患者氏名の確認)」といった比較的客観的で単純な評価項目については画像および音声から正解が判るため、評価者が正しく評価したかどうかの確認も可能である。実際、自己紹介を適切に行っているにもかかわらず「できなかった」と評価した例が2例、適切に患者氏名の確認を行ったが「できなかった」と評価した例が3例みられた。一方で、すでに考察したように一連の会話の流れのなかでの評価が必要な項目は「正解」を定めることは困難であり、特に医療面接課題ではこのような評価項目が入り混じっていることを理解したうえで合否判定を行うことが公平性を保つうえでも必要である。

31名の受験生を対象とした薬学OSCEトライアルの医療面接課題において得られた評価結果を統計的に解析し、また、課題の画像および音声を事後に精査し、受験生や模擬患者の発言内容を確認することで医療面接課題における評価方法の問題点について検討した。その結果、個々の評価項目ではその内容により評価者間の評価の一致度がばらつき、特に「身だしなみ」といった総合的な判断を要する項目や、「姿勢」「傾聴」といった評価者自身の主観が反映されやすい項目、あるいは一連の会話の中で他の項目と同時に出現する内容等に関して一致度が低いことがわかった。また、模擬患者からみた評価については、例数は少ないものの一般模擬患者と比べて教員による模擬患者では厳しい評価をする傾向がみられた。また、ビデオ映像を精査することでそれぞれの評価項目における問題点を抽出することができた。

OSCEにおける評価では、個別的項目評価による合否判定と概略評価による判定とが一致することが理想であるが、今回の検討結果から考えると、医療面接課題における個別的評価において信頼性・妥当性を高めるためには、特に一致度の低い項目に関して検討が必要である。一方、合否判定については面接技能・態度のマニュアル評価による手法の限界を認識し、現状では概略評価に重点を置くことが現実的と考える。

謝辞 今回の京都大学薬学部におけるOSCEトライアルの実施にあたりご支援頂いた、京都府薬剤師会、京都府病院薬剤師会、京都大学医学部附属病院薬剤部、京都大学模擬患者の会、京都薬科大学、摂南大学薬学部、同志社女子大学薬学部、京都大学薬学研究科の各位に深謝致します。

この論文内容の一部は、文部科学省平成18年度採択「地域医療等社会的ニーズに対応した質の高い医療人養成推進プログラム(医療GP)-先端医療の育・創薬を先導する薬剤師養成-」の支援によって行われました。

引用文献

- 1) 半谷真七子, 松葉和久, 松井俊和, 薬学生の臨床コミュニケーション教育の評価としての客観的臨床能力試験(OSCE)の試みとその評価, *医療薬学*, **31**, 606-619 (2005).
- 2) 高柳理早, 横山晴子, 林原絵美子, 成井浩二, 大関健志, 野口雅久, 安藤利亮, 山田安彦, 笹津備規, 薬学における客観的臨床能力試験(OSCE)の課題と評価設定に関する検討-医療薬学専攻大学院生を対象としたトライアルとその解析-, *薬学雑誌*, **126**, 83-91 (2006).
- 3) 向後麻里, 神山紀子, 根来孝治, 青木公子, 齋藤勲, 小林靖奈, 真下順一, 佐々木圭子, 戸部敏, 山元俊憲, 木内祐二, 佐藤均, 昭和大学薬学部で試行された客観的臨床能力試験(OSCE)における学生の達成率と評価内容の検討, *薬学雑誌*, **127**, 905-917 (2007).
- 4) 入江徹美, 薬学教育改革における共用試験の意義, *日本病院薬剤師会雑誌*, **43**, 1361-1364 (2007).
- 5) 斎藤百枝美, 中島康雄, 小佐野博史, 栗原順一, 渡邊真知子, 藤井靖史, 柳川幸重, 山岡桂子, 井上圭三, 帝京大学薬学部における小児科外来実習でのアンケートを用いた医療面接実習の有用性の評価, *医療薬学*, **33**, 741-747 (2007).
- 6) 鈴木栄一, 伊藤雅章, 青柳豊, 布施一郎, 田中恵子, 内藤眞, 山本正治, OSCEにおける評価の妥当性と信頼性についての検討-新潟大学医学部第1回OSCEを実施して-, *医学教育*, **34**, 37-44 (2003).
- 7) 福本陽平, 村上不二夫, 小早川節, 小野咲弥子, 村上泰昭, 田村周, 川崎勝, 医療面接における教員と模擬患者による学生評価について-山口大学医学部における3年間の検討-, *医学教育*, **35**, 229-234 (2004).
- 8) J.R. Landis, G.G. Koch, The measurement of observer agreement for categorical data., *Biometrics*, **33**, 159-174 (1977).