·临床教学·

客观结构化放射学考试在医学影像技术专业实习临床技能考核中的应用

张俊　刘佳怿　梁志鹏
南京医科大学附属逸夫医院医学影像科，南京　211100
通信作者：梁志鹏，Email：53113379@qq.com

【摘要】 目的　探讨客观结构化放射学考试（objective structured radiology examination，OSRE）在医学影像技术专业临床技能考核中应用价值，并对该评估工具进行可重复性评价。**方法**　选择南京医科大学附属逸夫医院医学影像科医学影像技术学专业实习生55人，随机分为11组（每组5人）；并设计9个不同的OSRE病例，每位学生均需完成3个OSRE病例；再由5位高年资主治医师或副主任医师对每个OSRE病例进行独立评分；实习结束后2个月，随机选择3个OSRE病例进行评分，以评估评分者之间的内部一致性和评估方式的可重复性。采用SPSS 16.0进行统计分析，使用多元线性回归分析学历与OSRE成绩的相关性。**结果**　所有学生的OSRE平均成绩为（12.0 ± 4.9）分。评估者间的ICC平均范围为0.36到0.80（P<0.001）；6个病例一致性良好，病例3的一致性较差，病例6和7的一致性中等；2个月后，4个评分者ICC等于或大于0.8（一致性极佳），评分者3的ICC为0.68，平均原始分数差的范围是−1~0.8。在多元线性回归中，学历的OR值为0.90（95%置信区间0.30~1.50，P=0.01），学历与OSRE评分呈正相关。**结论**　客观结构化放射学考试旨在通过使用预先设定的标准评分表，在标准化的临床环境中评估临床技能和知识，从而使临床知识能够以高水平和可重复的方式进行评估。

【关键词】 医学教育；　放射学；　客观结构化临床考试
【中图分类号】 R81
基金项目：南京医科大学康达学院2019年度教育研究课题（KD2019JYYJYB013、KD2019JYYJYB014）
DOI: 10.3760/cma.j.cn116021-20210322-01275

Application of objective structured radiology examination in clinical skills assessment of medical imaging technology

Zhang Jun, Liu Jiayi, Liang Zhipeng
Department of Radiology, Yifu Hospital Affiliated to Nanjing Medical University, Nanjing 211100, China
Corresponding author: Liang Zhipeng, Email: 53113379@qq.com

【Abstract】 Objective　To explore the application value of objective structured radiology examination (OSRE) in clinical skills assessment of medical imaging technology, and evaluate the repeatability of the assessment tool. **Methods**　A total of 55 interns majoring in medical imaging technology were randomly divided into 11 groups (5 in each group), and 9 different cases of OSRE were designed. Each student had to complete 3 cases of OSRE. Five senior attending doctors or deputy chief physicians independently evaluated each case of OSRE. Two months after the internship, three cases of OSRE were randomly selected for scoring to evaluate the internal consistency between raters and the repeatability of the evaluation method. SPSS 16.0 was used for statistical analysis, and multiple linear regression was used to analyze the correlation between academic qualifications and OSRE scores. **Results**　The OSRE average score of all the students was (12.0 ± 4.9), and the average ICC range between evaluators was 0.36 to 0.80 (P < 0.001). The consistency of 6 cases was good, that of case 3 was poor, and that of cases 6 and 7 was moderate. After 2 months, the ICC of 4 raters was equal to or greater than 0.8 (excellent consistency), and the ICC of rater 3 was 0.68, and the average original score difference ranged from −1 to 0.8. In multiple linear

regression, the OR value of education was 0.90 (95% confidence interval 0.30−1.50, $P = 0.01$), and education was positively correlated with OSRE score. **Conclusion** The OSRE aims to assess clinical skills and knowledge in a standardized clinical environment through the use of pre-set standard scoring scales, so that clinical knowledge can be evaluated in a high-level and repeatable manner.

【**Key words**】 Medical education; Radiology; Objective structured clinical examination

Fund program: 2019 Education and Research Project of Kangda College of Nanjing Medical University (KD2019JYYJYB013, KD2019JYYJYB014)

DOI: 10.3760/cma.j.cn116021-20210322-01275

客观结构化临床考试(objective structured clinical examination,OSCE)最初是由英国Harden教授于1975年提出的[1],作为评估医学生的临床理论与技能操作水平的一种考核方式。OSCE是在标准的临床环境中,通过预先设定的标准评分表来评估理论知识和技能水平,以便能够以高水平和可重复的方式来评估。越来越多医学教育研究已经证实OSCE是评估医学生和住院医师临床能力的有效方法[2],包括医学影像学教学。本研究尝试提出一种医学影像学教学所特有的OSCE方式。该评估方式以医学影像学教学为特点,客观结构化放射学考试(objective structured radiology examination,OSRE)是以客观结构化方式评估理论知识和技能水平,在不同的案例和不同的评价者之间具有可重复的结果,通过提供客观、结构化、可复制和低成本的方法,以类似于OSCE的形式来评估医学影像学教学,从而进一步提高教学质量。

1 对象与方法

1.1 研究对象

选取本科室医学影像技术学专业实习生55人,其中本科生32人、专科生23人;用简单随机法分为11组,每组5人。每位学生均需完成3个OSRE病例,9个OSRE病例按顺序给出,比如第1、4、7和10组的OSRE病例为1、2、3;第2、5、8和11组的OSRE病例为4、5、6;第3、6和9组的OSRE病例为7、8、9。本研究已获得机构审查委员会的批准,并且免除了获得每位学生同意的需要,并告知学生OSRE病例得分不计入实习结束时的最终成绩。每位学生在整个放射科轮换期间都接受标准课程的教学,但没有任何针对OSRE的相关教学。

1.2 OSRE设计

本研究设计9个医学影像学OSRE病例,每个病例均有5个问题(表1)。首先将这些OSRE病例对低年资住院医师进行初步测试,然后公开评价其成绩并对评分表进行修订,尽可能地纠正错误的评分。对于每个OSRE病例均设计一个计分表,其中包含对应于正确和错误计分的复选框,并对每一个正确或错误的评分都指定一个分值。每个OSRE病例包括1~3张的影像学图像,一个临床场景和结构化问题及一个标准化评分表。本研究设计5个问题来模拟临床医生可能以结构化方式执行的活动。如填写影像学检查方式所需的相关临床病史、识别临床的重要发现、形成总体印象及有关建议和随访的问题。所有病例图像均由本科室影像归档和通信系统(picture archiving and communication systems,PACS)提供。

1.3 OSRE病例

病例1为胸部正侧位显示右肺上野大叶性肺炎;病例2为腹部立位平片显示小肠梗阻;病例3包括2张硬膜下血肿不同层面的CT横断位图像;病例4为腕关节正侧位显示桡骨远端骨折;病例5为胸部后前位显示右侧胸腔积液;病例6为腹部仰卧位显示胃腔内鼻饲管;病例7为膝关节侧位显示髌上囊脂肪–积液平面;病例8为胸部正侧位显示左肺上野Ⅲ型肺结核;病例9为腹部立位平片显示消化道穿孔。

1.4 OSRE评分

OSRE评分表包含有正确和错误评分的特定分值,每个问题的分数在1到4分之间。例如,OSRE中的一个问题要求显示胸片的相关阳性和阴性结果,对正确回答给予正分,对错误回答给予负分。每个OSRE病例最高得分为20分、最低分只能是0。

1.5 评分方法

由5位高年资主治医师或副主任医师对每个OSRE病例进行独立评分。实习结束后2个月,随机选择3个OSRE病例进行评分,以评估评分者之间的内部一致性和评估方式的可重复性。本研究根据

表 1　医学影像学 OSRE 病例

病例	问题 1	问题 2	问题 3	问题 4	问题 5
1. 胸部正侧位	检查方法	病灶/病变部位	描述病灶	简单分析临床特点	疾病诊断
2. 腹部立位	检查方法	病灶/病变部位	描述病灶	简单分析临床特点	疾病诊断
3. 颅脑 CT 平扫	检查方法	病灶/病变部位	描述病灶	简单分析临床特点	疾病诊断
4. 腕关节正侧位	检查方法	病灶/病变部位	描述病灶	简单分析临床特点	疾病诊断
5. 胸部后前位	检查方法	病灶/病变部位	描述病灶	简单分析临床特点	疾病诊断
6. 腹部仰卧位	检查方法	病灶/病变部位	描述病灶	简单分析临床特点	疾病诊断
7. 膝关节正侧位	检查方法	病灶/病变部位	描述病灶	简单分析临床特点	疾病诊断
8. 胸部正侧位	检查方法	病灶/病变部位	描述病灶	简单分析临床特点	疾病诊断
9. 腹部立位	检查方法	病灶/病变部位	描述病灶	简单分析临床特点	疾病诊断

评分表设计了易于解释的测试和评分细则，因此没有针对学生的具体培训或指导。每个 OSRE 病例大约需要 30 s 到 1 min 来评分，并对学生的个性化体验形成反馈，以了解他们对 OSRE 病例的建议。学生对于 OSRE 病例和评分有任何疑问均可得到回答。

1.6　统计学方法

所有数据均采用 SPSS 16.0 统计软件进行分析。符合正态分布点计量资料用（均数 ± 标准差）表示，不符合正态分布采用中位数及范围表示。重测信度与评定者间信度采用组内相关系数（intraclass correlation coefficient，ICC）进行评价。使用 Landis 和 Koch 标准的 ICCs 分类系统，ICC 低于 0.4 为一致性较差，ICC 在 0.41 到 0.75 之间为一致性良好，ICC 大于 0.75 为一致性极佳。使用多元线性回归分析，OSRE 评分为结果变量、分组为解释变量、学历为控制变量，采用双边 t 检验比较本科生和专科生的 OSRE 平均成绩。检验水准 α=0.05。

2　结果

2.1　OSRE 评分

各组学生的 OSRE 成绩及各 OSRE 病例的平均成绩见表 2，学生的 OSRE 平均成绩为（12.0 ± 4.9）分。

表 2　OSRE 评分[$(\bar{x} \pm s)$，分]

分组	分数	病例	分数
1	13.0 ± 4.1	1	15.1 ± 4.1
2	8.3 ± 3.6	2	12.8 ± 4.5
3	12.9 ± 6.2	3	13.6 ± 3.5
4	15.1 ± 3.6	4	8.6 ± 4.0
5	11.5 ± 4.3	5	8.5 ± 3.2
6	12.6 ± 5.1	6	11.4 ± 4.3
7	13.3 ± 4.2	7	6.4 ± 3.2
8	9.2 ± 3.8	8	15.9 ± 4.0
9	12.2 ± 4.9	9	14.5 ± 3.0
10	14.0 ± 4.2		
11	9.0 ± 4.1		

2.2　重测信度分析

结果显示，评估者间的 ICC 平均范围为 0.36 到 0.80（P<0.001）（表 3）。在大多数情况下，9 个病例中有 6 个 ICC 大于 0.6（一致性良好），病例 3 的一致性较差，病例 6 和 7 的一致性中等。2 个月后，3 个随机 OSRE 病例（病例 1、2、3）的评分一致性显示，5 个评分者中有 4 个 ICC 等于或大于 0.8（一致性极佳），评分者 3 的 ICC 为 0.68，平均原始分数差的范围是 −1 到 0.8（表 4）。

表 3　具有组内相关系数的评分者之间的信度

OSRE	ICC	95% CI	P值
病例 1			
组 1	0.80	0.64~0.92	<0.001
组 4	0.71	0.52~0.87	<0.001
组 7	0.72	0.53~0.88	<0.001
组 10	0.65	0.41~0.86	<0.001
平均	0.72	N/A	N/A
病例 2			
组 1	0.79	0.64~0.90	<0.001
组 4	0.79	0.63~0.91	<0.001
组 7	0.90	0.81~0.96	<0.001
组 10	0.73	0.54~0.88	<0.001
平均	0.80	N/A	N/A
病例 3			
组 1	0.28	0.08~0.56	<0.001
组 4	0.31	0.10~0.60	<0.001
组 7	0.53	0.31~0.76	<0.001
组 10	0.32	0.11~0.61	<0.001
平均	0.36	N/A	N/A
病例 4			
组 2	0.76	0.58~0.89	<0.001
组 5	0.81	0.66~0.92	<0.001

续表

OSRE	ICC	95% CI	*P*值
组8	0.80	0.62~0.92	<0.001
组11	0.67	0.46~0.85	<0.001
平均	0.76	N/A	N/A
病例5			
组2	0.76	0.58~0.89	<0.001
组5	0.54	0.32~0.76	<0.001
组8	0.53	0.29~0.78	<0.001
组11	0.61	0.41~0.81	<0.001
平均	0.61	N/A	N/A
病例6			
组2	0.50	0.27~0.74	<0.001
组5	0.37	0.16~0.64	<0.001
组8	0.51	0.26~0.77	<0.001
组11	0.50	0.27~0.74	<0.001
平均	0.47	N/A	N/A
病例7			
组3	0.32	0.12~0.59	<0.001
组6	0.56	0.35~0.77	<0.001
组9	0.64	0.45~0.81	<0.001
平均	0.50	N/A	N/A
病例8			
组3	0.61	0.40~0.80	<0.001
组6	0.54	0.34~0.75	<0.001
组9	0.76	0.61~0.88	<0.001
平均	0.64	N/A	N/A
病例9			
组3	0.77	0.60~0.90	<0.001
组6	0.61	0.41~0.80	<0.001
组9	0.77	0.61~0.90	<0.001
平均	0.72	N/A	N/A

表4　2个月后各评分者可重复性研究

评分者	ICC	差异	*P*值
评分者1			
病例1	0.88	0.50	0.30
病例2	0.92	0.39	0.35
病例3	0.85	0.50	0.19
平均	0.88	0.46	N/A
评分者2			
病例1	0.89	– 1.34	0.02
病例2	0.85	– 2.06	< 0.001
病例3	0.73	0.25	0.70
平均	0.82	– 1.05	N/A
评分者3			
病例1	0.35	– 0.84	0.43
病例2	0.89	1.06	0.03
病例3	0.79	– 0.69	0.11
平均	0.68	– 0.16	N/A
评分者4			
病例1	0.95	0.06	0.83
病例2	0.82	0.50	0.37
病例3	0.78	– 0.44	0.39
平均	0.85	0.04	N/A
评分者5			
病例1	0.86	0.60	0.24
病例2	0.88	1.06	0.03
病例3	0.66	0.69	0.25
平均	0.80	0.78	N/A

表5　影响OSRE评分的多元线性回归分析

变量	*OR*值(95%CI)	*P*值
截距	9.3(7.4~11.2)	<0.001
分组	–0.02(–0.08~0.04)	0.480
学历	0.90(0.30~1.50)	0.010

2.3　多元线性回归分析

在多元线性回归中，分组的*OR*值为–0.02(95%置信区间为–0.08~0.04，*P*=0.480)，学历的*OR*值为0.90(95%置信区间为0.30~1.50，*P*=0.010)，学历与OSRE评分呈正相关(表5)。

3　讨论

目前，越来越多的方法被用来评估医学生在实习期间的表现，而有效且公平地评价成绩应该是主客观的结合。主观评价包括对学生履行职责的直接观察和书面评价或陈述；而客观评价包括考试(如笔试、网络考试)和报告书写等[3]。笔试是最常用的考试形式，尽管其评估专业知识水平的深度有限，但在教学工作中已广泛应用，在笔试中经常发现一些限制，如考生缺乏反馈及时效性差等。相反，口试可以评估更高水平的知识和理论，但受到评分不一致和潜在偏见等限制[4]。自20世纪70年代以来，OSCE逐渐广泛用于医学教育以评估医学生临床能力。OSCE可克服传统评估方法的不足，目前被公认为是用来评价医学生能力的一种可信的、有效的方法[5]。

本研究发现OSRE实现了客观、结构化、可复制和低成本的目标，大多数OSRE病例显示评估者的一致性良好(9个病例中有6个ICC大于0.6)。2个月后3个随机OSRE病例(病例1、2、3)的评分一致性显示，5个评分者中有4个ICC等于或大于0.8(一

致性极佳），学生表现出可重复的分数，并显示一致性极佳。评估不同OSRE病例显示，病例3、6和7的相关性较差至中等，而其他OSRE案例的ICC值都大于或等于0.6。但仍然认为OSRE是可重复的有效方法，因为对3个OSRE病例进行重新评分（相隔2个月）时的平均得分差异很小（–1~0.8）。本研究中，每个OSRE病例设置5个问题，这个数量既符合经济学节约原则，又符合OSCE考核模式良好信度要求。只是内外科专业考站设计的评分标准较为客观规范，而OSRE评分标准可能会由于考官的理解和评估缺少绝对客观的参考，故本研究相隔2个月后对OSRE进行重新评估。Morag等[6]通过对122名医学生进行研究发现，OSCE病例通过揭露个人缺陷提供了反馈的机会，学生能够轻松地回顾在不同临床症状（胸痛、腹痛等）及不同问题类型（检查方式）中的表现。本研究实施OSRE考核模式的优点是每周都有OSRE成绩反馈，这使得研究能够发现学生的不足及差距，并在期末反馈中向每个医学生提供更多的有效信息。据Agarwal等[7]研究发现，放射科应将OSCE作为其临床教学的重要组成部分，并解释理想的OSCE。具体地说，他们描述了一种具体的OSCE方法，该方法有10~20个工作站，每个工作站评估与放射科特定主题相关的活动[例如影像学读片或演示一种影像检查技术（如对患者进行腹部CT检查）]，5 min内完成一个工作站要求的任务，向考官演示或口头回答、笔试。他们的方法不同于本研究提出的方法，本研究OSRE是一个相对简单的考核方式且成本较低，与国内报道一致[8]。

尽管OSRE这种评估方式有很多优点，但也不缺乏局限性。一个局限是在所有OSRE病例中评估者之间的一致性并不显著，这可以通过针对OSRE的相关培训来弥补。另一局限性是本研究在单一机构进行且样本量较低，尚需大量样本与反馈，不断修改相关病例图像、问题表和评分表等。鉴于这些局限性，还可使用Internet来搭建OSRE云端。如果创建一个包含数百个OSRE案例的大型在线数据库，任何时间和任何地点均可随机下载一个案例用于医学生考试，可能会促使这种模式扩展到其他医学机构。

本研究中OSRE共9个病例，每个病例包括1~3张影像学图像、一个临床场景和结构化问题及一个标准化评分表，同时对医学生成绩进行客观、结构化和低成本的评估。结构化问题旨在评估实习生对理解临床病史和适应证、描述影像学表现、给出影像学诊断及提出下一步临床措施的能力。总之，OSRE旨在通过使用预先设定的标准评分表，在标准化的临床环境中评估临床技能和知识，从而使临床知识能够以高水平和可重复的方式进行评估。

利益冲突　所有作者声明无利益冲突

作者贡献声明　张俊、刘佳怿：提出论文构思、撰写论文；梁志鹏：总体把关、审定论文

参考文献

[1] Dwyer T, Glover Takahashi S, Kennedy Hynes M, et al. How to assess communication, professionalism, collaboration and the other intrinsic CanMEDS roles in orthopedic residents: use of an objective structured clinical examination (OSCE) [J]. Can J Surg, 2014, 57(4): 230-236. DOI: 10.1503/cjs.018813.

[2] van den Berk IAH, van de Ridder JMM, van Schaik JPJ. Radiology as part of an objective structured clinical examination on clinical skills [J]. Eur J Radiol, 2011, 78(3): 363-367. DOI: 10.1016/j.ejrad.2008.10.042.

[3] Zayyan M. Objective structured clinical examination: the assessment of choice [J]. Oman Med J, 2011, 26(4): 219-222. DOI: 10.5001/omj.2011.55.

[4] Aronowitz T, Aronowitz S, Mardin-Small J, et al. Using objective structured clinical examination (OSCE) as education in advanced practice registered nursing education [J]. J Prof Nurs, 2017, 33(2): 119-125. DOI: 10.1016/j.profnurs.2016.06.003.

[5] Emery AW, Rose-Innes E. Benefits of a peer-led mock-OSCE [J]. Med Teach, 2018, 40(3): 321-322. DOI: 10.1080/0142159X.2017.1392496.

[6] Morag E, Lieberman G, Volkan K, et al. Clinical competence assessment in radiology: introduction of an objective structured clinical examination in the medical school curriculum [J]. Acad Radiol, 2001, 8(1): 74-81. DOI: 10.1016/S1076-6332(03)80746-8.

[7] Agarwal A, Batra B, Sood A, et al. Objective structured clinical examination in radiology [J]. Indian J Radiol Imaging, 2010, 20(2): 83-88. DOI: 10.4103/0971-3026.63040.

[8] 姚隆，李华，曾闽榕，等. 基于器官系统的OSCE在临床实践教学中的应用[J]. 中华医学教育探索杂志，2020，19(4)：458-462. DOI: 10.3760/cma.j.cn116021-20190303-00105.

（收稿日期：2021-03-22）

（本文编辑：唐宗顺）