·诊断学教学改革专栏·

形成性评价在诊断学教学中的应用研究

岳荣铮[1]　张敬[2]　余婷[3]　曾锐[4]　左川[5]

[1]四川大学华西临床医学院/华西医院肾脏内科,成都　610041;[2]四川大学华西临床医学院/华西医院神经外科,成都　610041;[3]四川大学华西临床医学院/华西医院病理科,成都　610041;[4]四川大学华西临床医学院/华西医院党委办公室,成都　610041;[5]四川大学华西临床医学院/华西医院教务部,成都　610041

通信作者:左川,Email:zcsiren@163.com

【摘要】 目的　探讨形成性评价在诊断学教学中的策划构建和应用实施效果。**方法**　在诊断学本科教学中策划和制定形成性评价方案,并应用实施于教学实践。分析形成性评价成绩与期末考试成绩的相关性;并通过自制调查问卷和访谈形式,对实施效果进行调查分析。运用 SPSS 17.0 统计软件包进行统计学处理。期末考试成绩与形成性评价成绩比较采用线性相关分析,调查问卷项目评分比较采用 Wilcoxon 秩和检验。**结果**　在诊断学教学中形成性评价成绩与期末考试成绩呈显著正相关(r=0.559,P<0.001)。调查问卷结果显示,学生认为形成性评价方式明显优于终结性评价(Z=−8.632,P<0.001);同时学生普遍赞同形成性评价能够促进自主学习和自我监督,认为它能够对知识和技能的学习带来帮助。**结论**　在诊断学中应用形成性评价进行学业评估考核具备可行性,能够带来良好的学习效果,具有广阔的推广应用前景。

【关键词】 形成性评价;　终结性评价;　诊断学;　教学

【中图分类号】 R4

DOI: 10.3760/cma.j.cn116021-20200416-00443

Practice and exploration of formative assessment in diagnostics teaching

Yue Rongzheng[1], Zhang Jing[2], Yu Ting[3], Zeng Rui[4], Zuo Chuan[5]

[1]Department of Neurology, West China School of Medicine/West China Hospital, Sichuan University, Chengdu 610041, China; [2]Department of Neurosurgery, West China School of Medicine/West China Hospital, Sichuan University, Chengdu 610041, China; [3]Department of Pathology, West China School of Medicine / West China Hospital, Sichuan University, Chengdu 610041, China; [4]Party Committee Office, West China School of Medicine/West China Hospital, Sichuan University, Chengdu 610041, China; [5]Academic Affairs Division, West China School of Medicine / West China Hospital, Sichuan University, Chengdu 610041, China

Corresponding author: Zuo Chuan, Email: zcsiren@163.com

【Abstract】 Objective　To explore the method of building the formative assessment (FA) system in diagnostics teaching and investigate its implementation effect. **Methods**　We planned and formulated the formative assessment system in diagnostics teaching of undergraduate curriculum. The correlation between the scores of the formative assessment and final exam were analyzed. A questionnaire survey on students' attitudes towards FA and an interview on teachers' perspectives were conducted. SPSS 17.0 was used for statistical analysis. Linear correlation analysis was used to compare the final exam scores with formative assessment scores, and Wilcoxon rank sum test was used to compare the scores of questionnaire items. **Results**　The scores of FA were significantly positive related to the scores of final exam (r = 0.559, P < 0.001). The questionnaire analysis showed that students significantly valued the formative assessment

superior to the summative assessment ($Z=-8.632$, $P<0.001$). It was generally approved of the promotion of FA in students' self-direct learning and self-supervision and the advances in acquirement of knowledge and skills. **Conclusion** It is feasible of carrying FA in diagnostics teaching which could provide better learning outcome and might have an extensive prospect of application.

【Key words】 Formative assessment; Summative assessment; Diagnostics; Teaching

DOI: 10.3760/cma.j.cn116021-20200416-00443

形成性评价(formative assessment,FA),是指在一项活动中为获得更佳的活动效果而在活动过程中修正本身轨道所进行的评价[1]。美国芝加哥大学教育心理学家 Benjamin Bloom 将其引入教学领域,更具体地提出 FA 是指在教学过程中为了获得修改、完善所需的反馈信息而进行的评价[2]。形成性评价倡导改变传统的过分关注结果的终结性评价方式(summative assessment,SA),而通过在教学过程中对每个阶段和不同维度的学习进行不断的评价和反馈,形成更全面系统和客观可靠的评价;同时也赋予学生在教学中的责任感,使培养学生自主发展、自主性监控和调整能力成为可能,进而探索培养学生终身学习的态度、策略和能力[3]。

诊断学通常包括问诊、体格检查、症状学、正常和异常体征、实验室和辅助检查、病历书写、临床诊断思维等模块;各模块间既有独立性又相互关联。诊断学学习目标既强调对基本理论的掌握,也强调临床技能的运用,教学形式也相应丰富多样[4]。研究组尝试将形成性评价应用于诊断学教学中,希望以此为课程学习创建更客观公正和科学系统的评价体系。

1　对象与方法

1.1　研究对象

以整群抽样方法选取本校临床医学专业五年制 2014级学生 126 人纳入诊断学形成性评价体系,并对其学业能力进行考核评估。

1.2　诊断学课程形成性评价体系的创建

1.2.1　诊断学形成性评价的指标体系

形成性评价强调对学习过程的测评。其考核体系的构建和应用需强调对核心医学知识、技能、态度的评估;同时评估体系要能够反映学生学习中的进步情况、诊断知识难点和知识不足,能够激励学生自主学习的责任感[5]。四川大学组建诊断学课程专家团队,探讨形成性评价的指标体系,梳理、细化和提炼关键知识点和技能点,按照课程内容和进度安排,将课程考核细化为多层次、多时段、不同侧重面的阶段或过程评估(图 1)。专家组认为在诊断学教学中 FA 应包括临床理论知识掌握情况的阶段性评价、基本临床技能的分项测验评价、重要辅助检查及实验室检查的判读评价、临床综合能力评价(临床思维、分析决策能力、团队协作能力等)、平时表现评价、作业评价、学生自我及学生间同级评价等(表 1)。

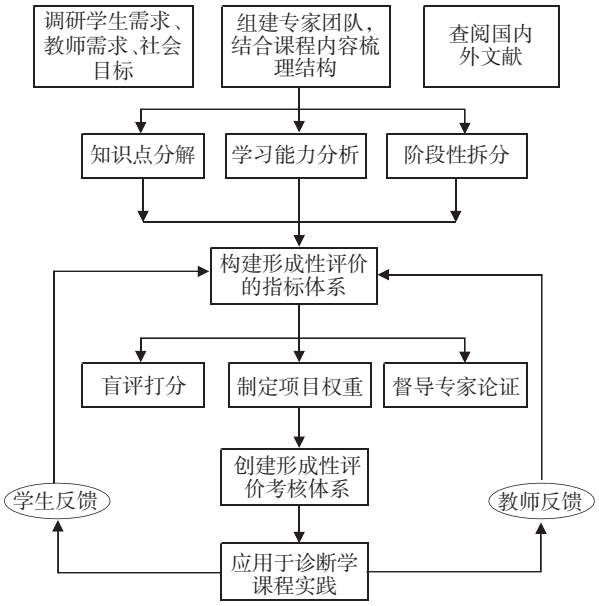


图 1　诊断学课程形成性评价学业考核体系建立模式图

1.2.2　成绩评定形式

学生的课程形成性评价成绩的计算方法为:

$$X=\sum_{i=1}^{n} X_i \times \lambda_i$$

式中,X 为课程形成性评价成绩;X_i 为第 i 项形成性评价考核的成绩;λ_i 为第 i 项考核成绩权重,即该项成绩占课程形成性评价成绩的百分比。课程形成性评价的成绩、各个单项成绩均按百分制记录。

1.2.3　反馈方式

FA 反馈方式包括书面反馈(书面作业、病历批

表 1　诊断学形成性评价的指标体系

评价构成	分值权重(%)	评价方式	评价目标	评价主体	评价时间
问诊	5	SP问诊考核	基本临床技能——问诊	SP教学指导员	第 8~12 周,第 19~23 周
体格检查	5	SP查体考核	基本临床技能——体格检查	SP教学指导员	第 8~12 周,第 19~23 周
症状学	5	TBL个人、小组及个人贡献成绩	诊断思维、团队协作、自主学习能力	教师、同级	症状学教学中穿插安排
体征	5	模拟人考核	临床技能	教师	第一学期末
心电图	5	PPT考试/机考/APP 考	辅助检查分析能力	教师	相应实验课结束后
实验诊断学	5	实验报告	实验室检查操作及分析能力	教师	相应实验课结束后
知识的提炼、应用及拓展	5	综述/视频/科普文作品	总结所学知识、撰写规范综述,或组成小组拍摄制作相关教学内容的小视频	教师、助教	第一学期末
阶段测验——网络测验	5	不定期网络测验(6~8 次)	阶段学习效果、主动的学习安排能力	助教	1~2 次/月
阶段测验——随堂测验	5	不定期随堂测验(6~8 次)	阶段学习效果、主动的学习安排能力	助教	1~2 次/月
教学反馈	5	网络调查问卷完成情况(课程架构设置/课堂/见习/SP)	学习过程的适应、思考和建议	学生	学生自己安排时间完成,要求每种反馈至少完成 1 次有效问卷
病历书写	5	病历书写(约 8 份病历)	医疗文书的书写能力	教师	每次床旁问诊查体后(7~8 次/月)
课堂表现	5	行课过程观察	学习态度和表现、职业素养等	教师	整个行课过程观察和评估

注:SP,即标准化病人(standardized patient)的缩写;TBL 为以团队为基础的学习(team-based learning)的缩写

改、综述审阅和交流、网络测验后的评分和解析等)、口头反馈(病历汇报、SP 考核的问题发现和及时反馈指导、TBL 教师反馈和组间讨论等)、示范反馈(视频演示临床操作过程、问诊查体示范等)。反馈过程注重反馈的及时性和相应的改进意见的提供。

1.3　评价方法

在诊断学教学过程中应用形成性评价对学生学业能力进行过程考核。方案实施过程中严格按照专家组制定的评价指标和原则,由各相关评价主体进行评分,每阶段评价后及时反馈。对学生期末考试成绩与形成性评价成绩进行相关分析。

自制《诊断学课程形成性评价调查问卷》调查学生对形成性评价的态度、教学效果等;同时了解学生对 FA 的建议和看法,以及对期末成绩与形成性评价成绩比例的权重看法。问卷通过网络平台发放和回收,包括 12 个条目,其中 10 个条目采用 Likert 5 级评价方式打分[6];1 个条目为开放性问题,采用填空形式采集信息;1 个条目为矩阵题,采集学生对成绩比例的权重。教师对形成性评价的态度采用访谈形式了解。

1.4　统计学方法

运用 SPSS 17.0 统计软件包进行统计学处理。资料用(均数 ± 标准差)或%表示,以 α=0.05 为检验水准。期末考试成绩与形成性评价成绩比较采用线性相关分析,调查问卷项目评分比较采用 Wilcoxon 秩和检验。

2　结果

2.1　成绩分析

对 126 名学生诊断学课程期末考试成绩(81.24 ± 9.02)与形成性评价成绩(90.52 ± 6.56)进行 Pearson 相关分析。各形成性评价分项成绩与期末考试成绩的相关性呈现类似趋势(r=0.466~0.661,P<0.001),显示学生期末考试成绩和形成性评价成绩呈中等强度正相关,提示形成性评价成绩的提高能够正向影响期末考试成绩。

2.2　问卷调查结果分析

发放调查问卷 126 份,收回有效问卷 126 份,有效率 100%。结果显示:大部分学生对既往的终结性评价考核方式持负面评价;绝大多数学生对现行的诊断学课程形成性评价持满意态度,并支持这种考核方式的进一步推广应用。对既往传统终结性考核方式(2.81 ± 0.93)和现行形成性评价方式态度(4.11 ± 0.82)的比较,提示学生对这两种考核评价方式的态度有显著差别。在“FA 能促进我的自主学习”“FA 能促进我自我监督和不断改进”“FA 对理论知识掌握有帮助”“FA 对临床技能掌握有帮助”和“FA 对培养持续学习和终身学习的习惯有帮助”等方面,学生评分均超过 4 分(表 2)。此外,在开放性问题“你对现行诊断学形成性评价方式的建议和意见”的回答中,多数学生认为现行评价方式已经比较完善了,多元化考核设计能够帮助学生提高学习能力;部分学生提

出可尝试将病历书写等作业批阅部分改成由同学间相互批改和评阅,以利于相互交流和促进;个别学生提到网络测验和随堂测验次数偏多。在课程负荷方面,学生普遍赞同“FA 减少了期末前的复习压力”,同时大部分学生也认为“FA 增加了平时的课业负担”。在矩阵题“你认为课程的过程评估考核与期末考试所占的比例应该如何”中,学生认为形成性评价分值比例应为(61.41 ± 13.23)%,期末测验分值应为(38.59 ± 13.23)%,与本次专家组拟定的权重比例基本一致。

表 2　诊断学课程调查问卷结果分析

项目	得分	选择情况[n(%)]
1. 你对既往终结性评价考核的态度	2.81 ± 0.93	5(3.94)
2. 你对诊断学现行的形成性评价方式的态度	4.11 ± 0.82	45(35.43)
3. FA 能促进我的自主学习	4.45 ± 0.61	65(51.18)
4. FA 能促进我自我监督和不断改进	4.23 ± 0.69	47(37.01)
5. FA 对理论知识掌握有帮助	4.17 ± 0.73	46(36.22)
6. FA 对临床技能掌握有帮助	4.26 ± 0.65	47(37.01)
7. FA 对培养持续学习和终身学习的习惯有帮助	4.18 ± 0.75	48(37.8)
8. FA 增加了平时的课业负担	3.77 ± 0.9	25(19.69)
9. FA 减少了期末前的复习压力	4.22 ± 0.8	54(42.52)
10. 希望后面教学中应用 FA 模式	4.37 ± 0.73	65(51.18)

2.3　教师对 FA 的态度

对教师的访谈发现,应用 FA 过程中,教师教学投入精力大大增加。教师需要较以往更多地关注学生的学习过程和个体发展,并不断给予群体或个体化反馈指导;同时也会花费更多的精力了解学生的反馈意见,以及时调整自己的教学节奏、重点和教学方式方法,教学的压力和挑战有显著增加。但是,尽管如此,教师对于更热烈活跃的课堂氛围和学生学习热情的高涨仍然印象深刻,对 FA 评价模式普遍持正面评价,并希望在今后的教学中进一步推广。同时,教师普遍认为,形成性评价的考核方式成功开展的关键因素是构建科学合理的考评体系,并确保反馈的及时性,以指导学生明确下一步努力的方向和目标,教师及时有效调整教学节奏。

3　讨论

本研究结果显示,对于在诊断学课程中应用形成性评价,教师和学生普遍持正面评价。学生普遍认为,FA 在促进学生自主学习、自我监督、培养良好学习习惯等方面有价值,且能调整课业压力的分布。成绩分析显示,形成性评价与期末考试成绩呈显著正相关。

本研究对学生期末考试成绩和形成性评价成绩的相关分析显示,FA 与期末考试成绩呈中等程度显著正相关,形成性评价各分项成绩与期末考试成绩的相关性也呈现类似趋势,并无任何分项成绩显示出与期末考试呈现更突出的相关性。可能由于本课程形成性评价各分项均考虑并设置了多阶段多次的考核和反馈,使得每项考核过程中学生均可及时发现自身问题并调整和加强跟进。这提示通过促进过程学习、提高形成性评价成绩能够相应提高期末考试成绩。形成性评价过程中学生通过分阶段的测试和及时反馈,可以不断地调整学习方向和学习方法。这种积累能够带来更好的期末成绩,可能是形成性评价成绩与期末成绩相关性的解释。

同时调查结果显示,在“FA 能促进我的自主学习、自我监督和不断改进”“FA 对理论知识、临床技能掌握有帮助”和“FA 对培养持续学习和终身学习的习惯有帮助”等方面,学生普遍持赞成态度;教师访谈也发现,教师对 FA 持认可、赞成态度。

形成性评价相对于传统终结性评价的特点和优势在于:首先,FA 注重调动和发展学生自主学习能力。通过阶段性测试评估,促进学生主动安排好自己的学习活动。FA 最基本的目的就是确认学生目前的学习水平和预计达到的教育目标之间的差距,而通过科学构建的 FA 评估系统能够更合理地设置这种差距的大小,以遵循学生认知模式,使学生依据个体化的成就和差距,调整学习重点和学习节奏,从而达到自主学习和自我监控的目标[7-9]。本研究调查问卷结果显示,在“FA 能促进我的自主学习”“FA 能促进我自我监督和不断改进”和“FA 对培养持续学习和终身学习的习惯有帮助”等方面,学生均给予了很高的评分,也印证了 FA 的这种优势。其次,FA 注重学生知识掌握、实践操作能力和创新能力等综合素质的提高[2-3]。本研究问卷中,“FA 对理论知识掌握有帮助”“FA 对临床技能掌握有帮助”两项学生评分均超过 4 分。这显示学生在 FA 评价体系指导下能够切身体会到这种评价方式对学习的促进作用。通过将课程内容分解成不同维度的测试,引导学生对每项技能、知识以及职业素质投入更多的关注和自觉训练,可能是带来良好学习收获的原因。既往国外研究认为 FA 并不能如预期那样很好地促进学生投入学习,也并未发现采用形成性评价能够提高期末终结性考核的成绩[10-12]。本研究未设置平行对照组进行研究,但调查问卷显示学生普遍认为形成性评价

能够促进学习投入度和提高学业成就。这可能是由于既往研究中过程性评价仅作为学生自我督促和自我测试的自觉行为,并未真正纳入考核体系,而本研究则将多阶段的形成性评价均纳入学业考核体系,在外部动机上给予了学生足够的激励。再次,FA 注重评价的过程,而不是评价的结果。FA 的分阶段评估模式引导学生在课程运行的整个阶段不断累积知识技能,一改以往期末考试前突击应付的局面。本研究对学生课业负荷方面的调查结果显示,学生在“FA 减少了期末前的复习压力” 项目上给予了 4分以上的高分赞同评价,同时也在“FA 增加了平时的课业负担”上给予了 3 分以上的评分。这显示 FA 方式的确扭转了既往平日松散而考前突击复习的局面,分散了课业压力,引导学生把精力更多地投入平时的学习和积累。最后,FA 是具备自我完善机制的评价系统[2-3]。从本研究的实践中可见,FA 通过多样和及时性的反馈,让学生在不断的测评、反馈、修正或改进过程中趋向于进步和完善;同时也促进学生参与到教学主体中来,通过对课程内容设置、教学方式安排及教师能力的评价反馈,使课程本身和教师教学能力均不断修正和改善。FA 的反馈机制调动了教与学的两个主体,是确保教学质量的关键。

本研究调查问卷在矩阵题“你认为课程的过程评估考核与期末考试所占的比例应该如何划分”中,学生对形成性评价和期末测验分值比例的普遍意见与专家团队拟定的权重比例一致。这提示诊断学形成性评价考核体系构建基本合理。教师访谈和学生反馈表明,要确保形成性评价方式实施效果,应做到以下几点。首先是专家团队细致总结教学过程中的可提取和设定的考核阶段和考核维度,创建基本合理的考评构架。其次是确保评价结果反馈的多样性和及时性。反馈多样性可以使师生从不同层面获得下一步行为的合理建议;信息及时反馈可以及时引导和鼓励学生,并调整其学习方式和节奏,最后也能让教师在反馈中不断优化完善自己的教学模式和细节[13-14]。最后,还需要意识到,在临床教育教学环境中使用形成性评估方法并不是一种独立的存在,而是受到无数因素的影响,如各种人际关系、环境因素、教育传统、情感问题和各种主客观的微妙的抵抗形式等。如何看待及解决这些问题,使形成性评估方法在医学教育中得到适当的实施,是值得深思和面对的挑战[15]。

本研究的优势和局限性:本研究在内容涵盖庞大繁杂的诊断学教学中开展如此细化的形成性评价,是教学实践中的首创,具备良好的实践指导价值;但鉴于考核公平性和教育伦理学问题,本研究设计未设立平行对照组,使研究结果的推广有一定的局限性。

利益冲突　所有作者均声明不存在利益冲突

作者贡献声明　岳荣铮:撰写初始文稿、协调研究进度;张敬、余婷:项目执行和推进、国内外研究评价;曾锐:文稿修订和审理;左川:提出研究思路、协调研究进度、审理最终文稿

参考文献

[1] Scriven MS. Perspectives of curriculum evaluation [M]. Chicago: Rand McNally, 1967: 121-126.

[2] Bloom BS, Hastings JT, Madaus GF. Handbook on formative and summative evaluation of student learning [M]. New York: McGraw-Hill Book Company, 1971: 312-331.

[3] Torrance H. Formative assessment: some theoretical problems and empirical questions [J]. Cambridge J Educ, 1993, 23(3): 333-343. DOI: 10.1080/0305764930230310.

[4] 左川, 曾静, 卿平, 等. 互动参与式教学在诊断学教学中的应用[J]. 华西医学, 2012, 27(7): 1091-1094.

[5] Geraldine M. Quality assurance in medical education: a practical guide [M]. Springer, 2013: 39-41.

[6] Likert, RA. A technique for the measurement of attitudes [J]. Archieves of Psychology, 1932(140): 1-55.

[7] Sadler DR. Formative assessment and the design of instructional systems [J]. Instructional Science, 1989, 18(2): 119-144. DOI: 10.1007/BF00117714.

[8] Kämmer JE, Hautz WE, März M. Self-monitoring accuracy does not increase throughout undergraduate medical education [J]. Med Educ, 2020, 54(4): 320-327. DOI: 10.1111/medu.14057.

[9] Perry M, Linn A, Munzer BW, et al. Programmatic assessment in emergency medicine: implementation of best practices [J]. J Grad Med Educ, 2018, 10(1): 84-90. DOI: 10.4300/JGME-D-17-00094.1.

[10] Palmer E, Devitt P. The assessment of a structured online formative assessment program: a andomized controlled trial [J]. BMC Med Educ, 2014(14): 8. DOI: 10.1186/1472-6920-14-8.

[11] Raupach T, Brown J, Anders S, et al. Summative assessments are more powerful drivers of student learning than resource intensive teaching formats [J]. BMC Med, 2013 (11): 61. DOI: 10.1186/1741-7015-11-61.

[12] Palmer EJ, Devitt PG. Limitations of student-driven formative assessment in a clinical clerkship. a andomized controlled trial [J]. BMC Med Educ, 2008(8): 29. DOI: 10.1186/1472-6920-8-29.

[13] Bok HGJ, Teunissen PW, Favier RP, et al. Programmatic assessment of competency-basedworkplace learning: when theory meets practice [J]. BMC Med Educ, 2013(13): 123. DOI: 10.1186/1472-6920-13-123.

[14] de Bruin ABH, Dunlosky J, Cavalcanti RB. Monitoring and regulation of learning in medical education: the need for predictive cues [J]. Med Educ, 2017, 51(6): 575-584. DOI: 10.1111/medu.13267.

[15] Andreassen P, Malling B. How are formative assessment methods used in the clinical setting? a qualitative study [J]. Int J Med Educ, 2019(10): 208-215. DOI: 10.5116/ijme.5db3.62e3.

（收稿日期:2020-04-16）
（本文编辑:唐宗顺）