

Original Paper

# Evidence of Construct Validity of Computer-Based Tests for Clinical Reasoning: Instrument Validation Study

Tianming Zuo<sup>1</sup>, MD, PhD; Baozhi Sun<sup>1</sup>, MD, PhD; Xu Guan<sup>2</sup>, MD; Bin Zheng<sup>3</sup>, MD, PhD; Bo Qu<sup>1</sup>, MD, PhD

<sup>1</sup>Institute for International Health Professions Education and Research, China Medical University, Shenyang, China

<sup>2</sup>Education Center for Clinical Skills Practice, China Medical University, Shenyang, China

<sup>3</sup>Surgical Simulation Research Lab, Department of Surgery, University of Alberta, Edmonton, AB, Canada

**Corresponding Author:**

Bo Qu, MD, PhD

Institute for International Health Professions Education and Research

China Medical University

No. 77 Puhe Road

Shenyang, 110122

China

Phone: 86 189 0091 0198

Email: [qubo6666@163.com](mailto:qubo6666@163.com)

## Abstract

**Background:** Clinical reasoning (CR) is a fundamental skill for all medical students. In our medical education system, however, there are shortcomings in the conventional methods of teaching CR. New technology is needed to enhance our CR teaching, especially as we are facing an influx of new health trainees. China Medical University (CMU), in response to this need, has developed a computer-based CR training system (CMU-CBCRT).

**Objective:** We aimed to find evidence of construct validity of the CMU-CBCRT.

**Methods:** We recruited 385 students from fifth year undergraduates to postgraduate year (PGY) 3 to complete the test on CMU-CBCRT. The known-groups technique was used to evaluate the construct validity of the CBCRT by comparing the test scores among 4 training levels (fifth year MD, PGY-1, PGY-2, and PGY-3).

**Results:** We found that test scores increased with years of training. Significant differences were found in the test scores on information collection, diagnosis, and treatment and total scores among different training years of participants. However, significant results were not found for treatment errors.

**Conclusions:** We provided evidence of construct validity of the CMU-CBCRT, which could determine the CR skills of medical students at varying early stage in their careers.

(*JMIR Serious Games* 2021;9(4):e17670) doi: [10.2196/17670](https://doi.org/10.2196/17670)

**KEYWORDS**

medical education; assessment; computer-based test; clinical reasoning; validity

## Introduction

Each year, several hundred thousand students enter medical school, all of whom need to equip themselves with the necessary health care skills and knowledge [1]. Since 2014, the vast majority of Chinese medical students attend a 5-year program after high school to earn a bachelor's degree. Then, they work in a 1-year clinical internship before taking the nation's standardized medical licensure exams. If successful, they may register as medical practitioners. Postgraduate training in medical specialties is standardized to 3-year programs with the final credential called Master of Medicine; this is now required

of all clinical practitioners. In addition to learning a broad range of medical knowledge and practicing dexterity in hands, practitioners need to learn how to collect information from patients, process this information, and make accurate diagnostic decisions, similar to the expectations from a senior physician [2,3]. Clinical reasoning (CR) is a fundamental skill that separates medical personnel from other professionals. William Osler, a legendary pioneer medical educator, emphasized proper physical examination and diagnostic reasoning while maintaining the intimate physician-patient relationship. His teachings have resonated with generations of physicians [4]. Strictly speaking, CR refers to the procedure of collecting and

integrating patient information from various sources to arrive at a diagnosis and management plan; it is usually case specific [5]. Every medical teaching institute makes a great effort to understand the nature of CR and improve strategies for teaching CR skills to health trainees [6]. However, the conventional methods that are used in our education system today are not optimal [7-11].

Traditionally, CR is taught in the classroom (didactic lecture) and by the patient's side (clinical clerkship) [12-15]. A recent focus of integrating problem-based learning (PBL) has significantly improved the quality of CR education [6,12,16-18]. However, PBL relies heavily on the involvement and commitment of faculty instructors, which may not always be feasible [16,19]. Fidelity of case is also a problem compared to patient-side education [12]. Acquiring patient information by reading PBL cases from charts is quite a different experience than taking information directly from patients. Although instructors are making PBL cases in collaboration with clinicians, students still report a lack of case variety [17,20]. Creating sufficient clinical cases with clinical fidelity for CR training is a difficult task. Due to the above reasons, new technology is needed to improve our CR teaching.

In contrast to a paper- or lecture-based curriculum, computer-based CR training allows trainees to interactively take information from patients in a step-by-step process. There is also the possibility of accumulating a large volume of cases through international collaboration.

Currently, computer-based CR training can have different interfaces such as text, graphics, and animation [21]. The text-based CR training system is most widely used [22]. It is easy to create from clinical cases and deliver in the format of multiple choice questions or direct interface [23]. While medical images (including x-ray films, electrocardiograms, photos of lesions, etc) are required to give students more clinical information, graphic interface is also necessary. In several graphical models, illustrations of patients (in drawing or 2D pictures) can be used to create interactive experience for students when they collect information from patients [24]. Some computer-based CR training includes 3D animation or virtual reality technology to simulate the clinical scenario with high fidelity. However, the cost of creating 3D animation and virtual reality scenarios is much higher than the other computer-based CR models. It is difficult to create virtual patients without a team of technicians and instructional designers (Table 1).

**Table 1.** Types of computer-based clinical reasoning simulations and comparison.

Media	Advantage	Disadvantage
Text based	Relatively easy and rapid to develop; less expensive	Low level of fidelity
Graphic and animation based	Presents rich clinical evidence; moderate cost with enhanced fidelity	Replicates only part of clinical settings; low level of interactivity
Virtual reality	Combines highly sophisticated, life-like models with computer animations; can provide interactivity and feedback	Challenge to developers; often expensive

Sponsored by the National Medical Examination Center of China, China Medical University (CMU) started to develop computer-based CR training system in 2001. Educators and researchers at the Institute for International Health Professions Education and Research of CMU began to work with clinicians to develop cases for training CR skills and established the computer-based CR testing (CBCRT) system. Since 2002, CBCRT has been used as one part in the final comprehensive examinations of CMU to test the clinical skills of undergraduate students.

The CBCRT is composed of 5 interactive modules that allow students to interact with simulations to complete tasks: (1) history taking and physical examination, (2) writing orders and obtaining lab and medical imaging results, (3) reviewing obtained results, (4) working out diagnosis and differential diagnosis, and (5) observing the patient's condition change at different phases and changing locations for managing different therapies. The main features of the CMU-CBCRT virtual patient are displayed in Figure 1.

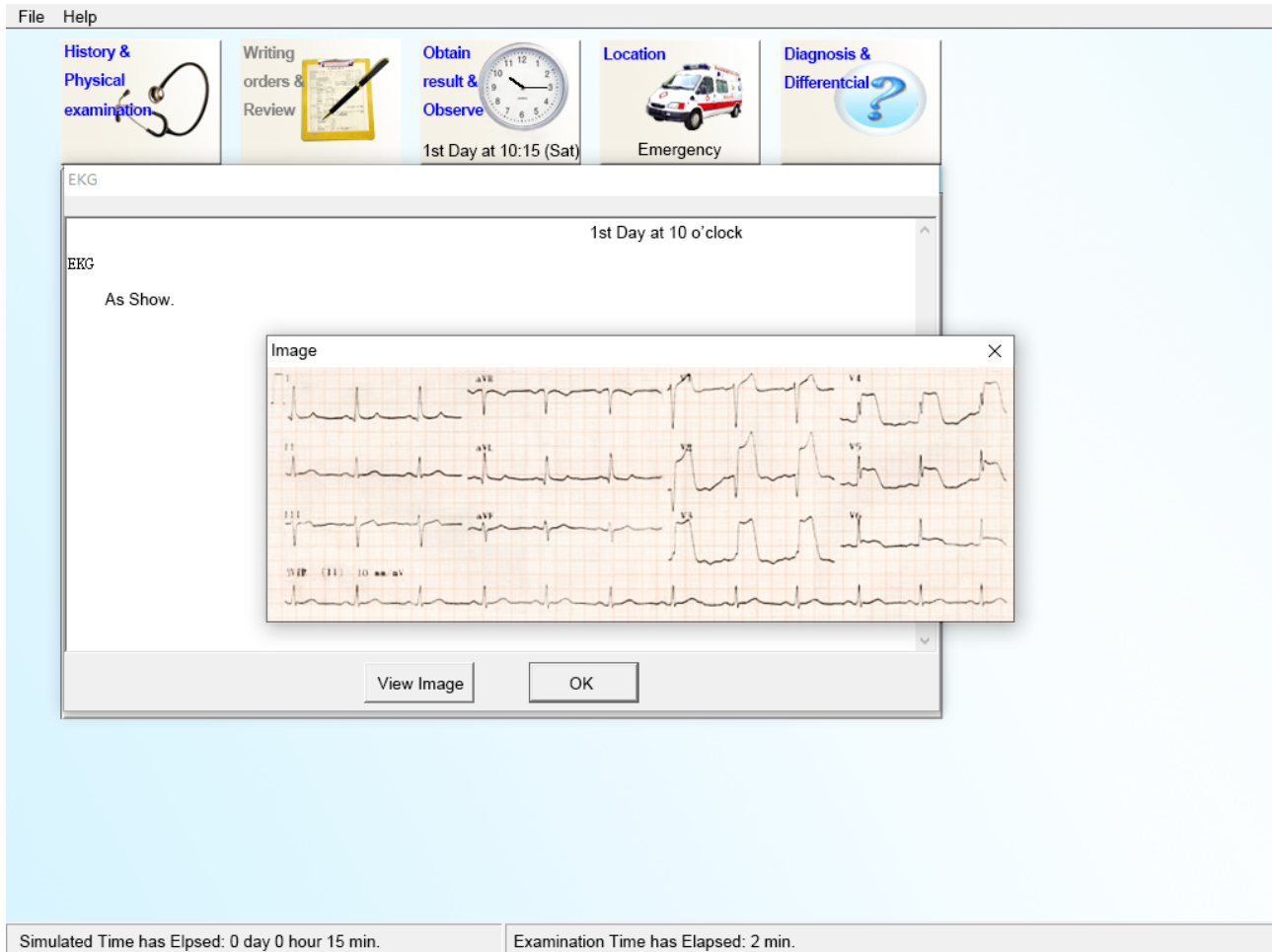
To test the face validity of the CMU-CBCRT, we called a series of meetings with physicians and surgeons at which we screened and selected key information on each clinical topic for CR training. When a CR case was developed, our clinical team was surveyed to verify their clinical relevance. They then evaluated the interactive interface and rated their level of satisfaction.

To briefly summarize, the CBCRT provided clinical features of patients including history and physical and laboratory findings and then requires students to make a diagnosis as well as a treatment plan for the simulated patients. The CBCRT has also been welcomed by the examinees based on their positive feedback toward the system. Of 300 students surveyed using the questionnaire, 99.4% enjoyed participating in the CBCRT examination; 95.9% believed that the system accurately represents the real clinical environment; 72.5% agreed that the CBCRT is a better tool for teaching their clinical abilities. We can thus believe that the face validity of the CBCRT is satisfactory.

However, face validity is the weakest form of validity evidence. It can be only used at the primary stage of designing an assessment method [25]. We need to look into the structure of the CBCRT in detail to find more evidence of its validity, especially since there is a paucity of validity evidence for computer-based CR training [22]. In China, this study is the first of its type.

This study investigates construct validity of the CMU-CBCRT in medical trainees over 5 medical school. We hypothesize that the CMU-CBCRT will be able to determine CR level among different years of health trainees; specifically, senior trainees will achieve higher CBCRT scores than juniors.

**Figure 1.** Screenshot displaying the main features of the China Medical University–computer-based clinical reasoning testing.



## Methods

### Ethical Statement

Methods used for the project were reviewed and approved by the ethical review boards of the CMU (ERB 2016-027) and the 5 medical schools. Informed consent was obtained from each participant before they started the test with the CMU-CBCRT.

### Testing Sites

From November 24 to December 8, 2016, we implemented the CMU-CBCRT system in 5 collaborative medical schools: China

Medical University, Fudan University School of Medicine, Sun Yat-sen University School of Medicine, Xuzhou Medical University, and Binzhou Medical College.

### Students

In China, medical students start their clerkships on the fifth year of medical training. The clinical training will continue for their 3 postgraduate years (PGYs). PGY 1 to PGY 3 is similar to the residency in North American medical school. We recruited students from fifth year undergraduates to PGY 3. The actual number of participants from each of 5 medical schools and their training years are shown as [Table 2](#).

**Table 2.** Students from the 5 medical schools and their training years.

Schools	Fifth year medical student	PGY-1 <sup>a</sup>	PGY-2	PGY-3	Subtotal
China Medical University	40	18	7	2	67
Fudan University School of Medicine	17	41	16	18	92
Sun Yat-sen University School of Medicine	12	28	20	12	72
Xuzhou Medical University	20	19	20	21	80
Binzhou Medical College	20	19	19	16	74
Total	109	125	82	69	385

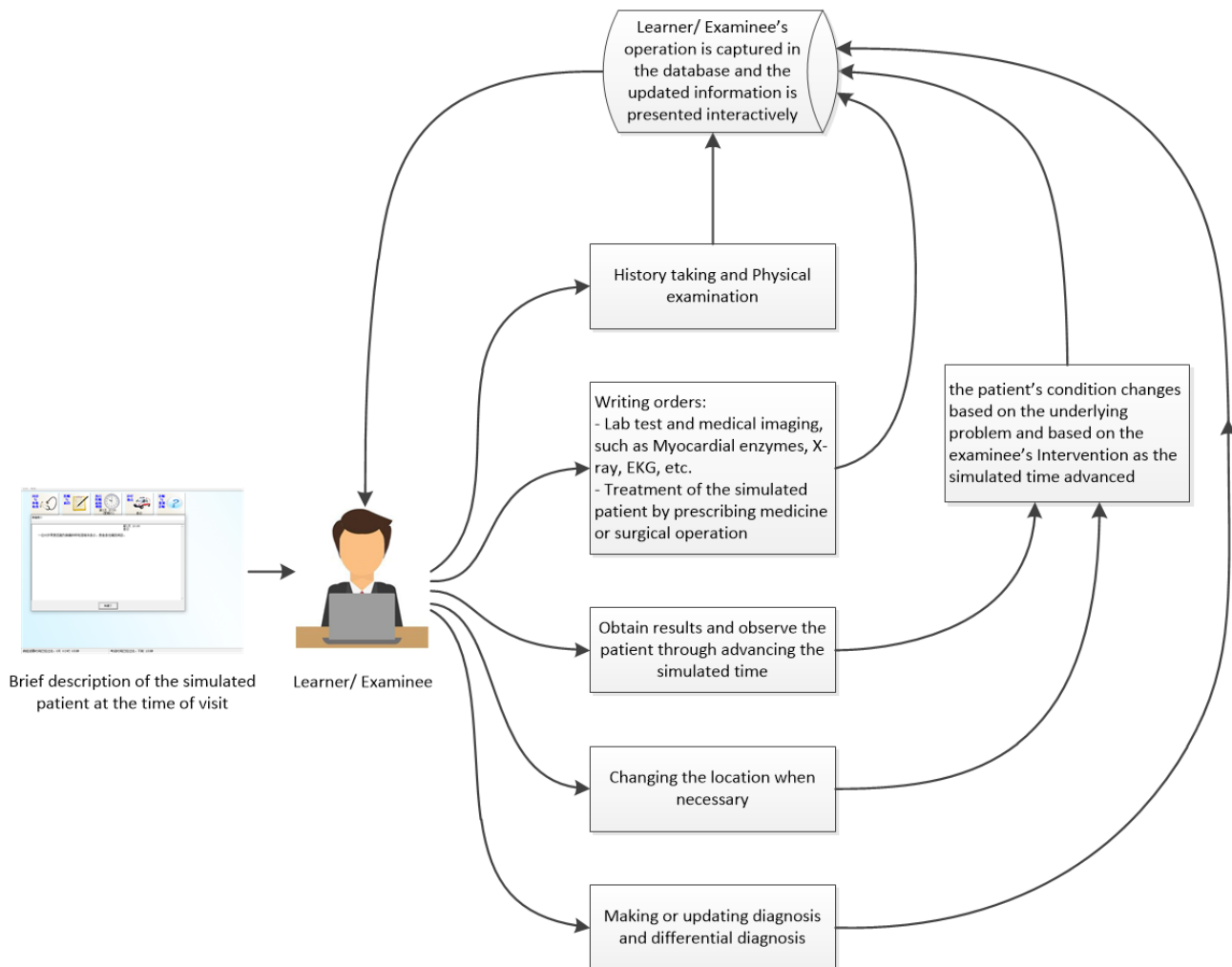
<sup>a</sup>PGY: postgraduate year.

## Measures

Before testing, each student was asked to watch a 5-minute presentation and get familiar with the testing interface. Demographics and level of medical training were surveyed and recorded. The computer recorded participants' typing and computer activity, including the typing and performance times. The interaction between a learner and how data are captured is

displayed in Figure 2. Once completing the testing on CMU-CBCRT, the system calculated and recorded their total score by comparing the participants' transaction list with the scoring scheme defined by the case developers committee (Multimedia Appendix 1). Subscores on these 4 different areas: information collection, diagnosis, treatment, and treatment error are computed and recorded as well (Multimedia Appendix 2).

**Figure 2.** Outline of interaction flow through China Medical University–computer-based clinical reasoning testing.



## Statistical Model

The known-groups technique was used to evaluate the construct validity of the CBCRT by comparing the scores among the fifth year MD, PGY-1, PGY-2, and PGY-3 participants. Testing scores, including total and subtotal, were compared over the 4 training groups using a 1-way analysis of variance (ANOVA). Results were reported as mean and standard deviation.  $P \leq .50$  was considered a significant difference among testing groups.

## Results

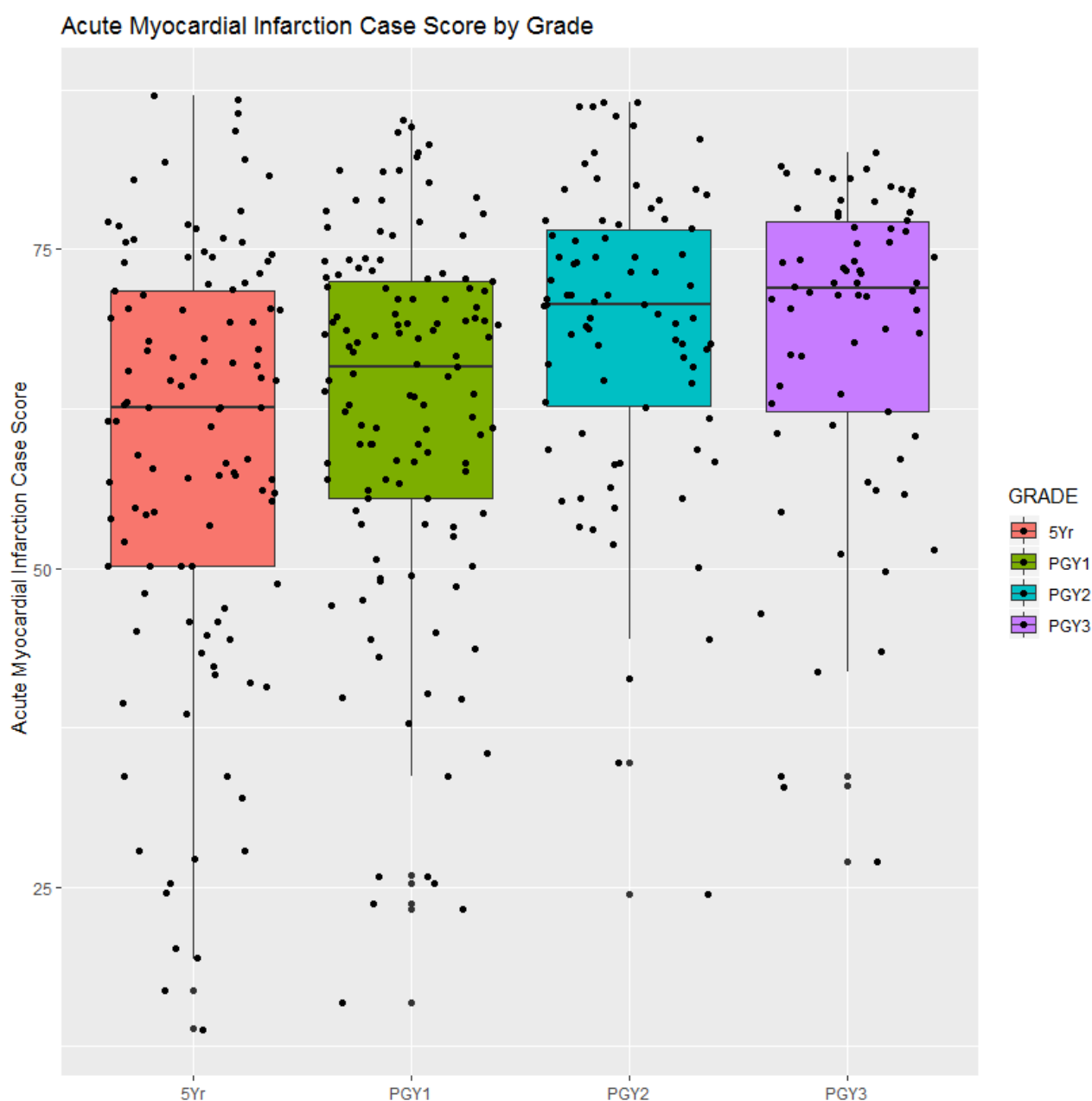
### Total Score

ANOVA revealed a group difference in total score among training levels ( $P < .001$ ). As shown in Table 3 and Figure 3, the score of the fifth year MD students (59.01 [SD 16.68]) was significantly lower than the PGY-2 (68.68 [SD 11.76]) and PGY-3 (68.06 [SD 12.67]) students; the total score of PGY-1 students was also significantly lower than the PGY-2 and PGY-3 students.

**Table 3.** Students from the 5 medical schools and their training years.

Scores	Fifth year medical student, mean (SD)	PGY-1 <sup>a</sup> , mean (SD)	PGY-2, mean (SD)	PGY-3, mean (SD)	P value
Information collection	43.42 (12.63)	46.70 (11.48)	49.73 (9.12)	51.38 (9.08)	<.001
Diagnosis	10.90 (4.74)	11.24 (4.97)	12.76 (3.90)	11.25 (4.22)	.034
Treatment	4.79 (3.81)	4.61 (3.36)	6.19 (3.73)	5.45 (3.72)	.013
Treatment error	-0.06 (0.23)	-0.04 (0.20)	0.00 (0.00)	-0.01 (0.12)	.13
Total	59.01 (16.68)	62.50 (14.45)	68.68 (11.76)	68.06 (12.67)	.001

<sup>a</sup>PGY: postgraduate year.

**Figure 3.** Total score of students over training years.

### Subscore

ANOVA revealed group differences by training level between information collection ( $P<.001$ ), diagnosis ( $P=.03$ ), and

treatment ( $P=.01$ ) scores, but not on treatment error ( $P=.13$ ) score. As shown in Figure 4, the information collection scores of the fifth year MD students (43.42 [SD 12.63]) were significantly lower than the PGY-1 (46.70 [SD 11.48]), PGY-2

(49.73 [SD 9.12]), and PGY-3 (51.38 [SD 9.08]) students; information collection scores of PGY-1 students were also significantly lower than the PGY-3 students. As shown in Figure 5, the diagnosis scores of the fifth year MD (10.90 [SD 4.74]), PGY-1 (11.24 [SD 4.97]), and PGY-3 (11.25 [SD 4.22]) students

were significantly lower than the PGY-2 (12.76 [SD 3.90]) students. As shown in Figure 6, treatment scores of the fifth year MD (4.79 [SD 3.81]) and PGY-1 (4.61 [SD 3.36]) students were significantly lower than the PGY-2 (6.19 [SD 3.73]) students.

Figure 4. Subscore for information collection of students over training years.

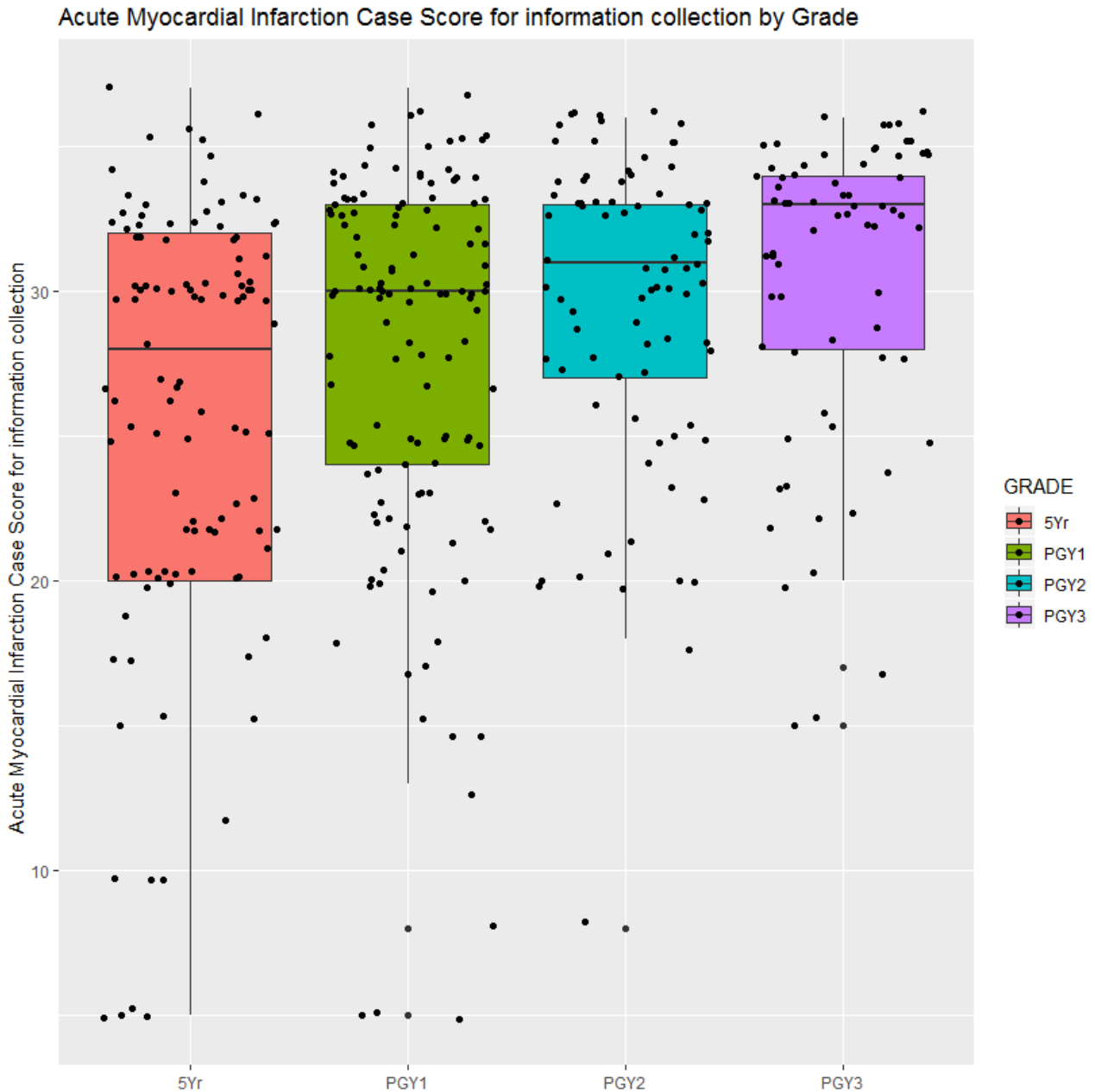
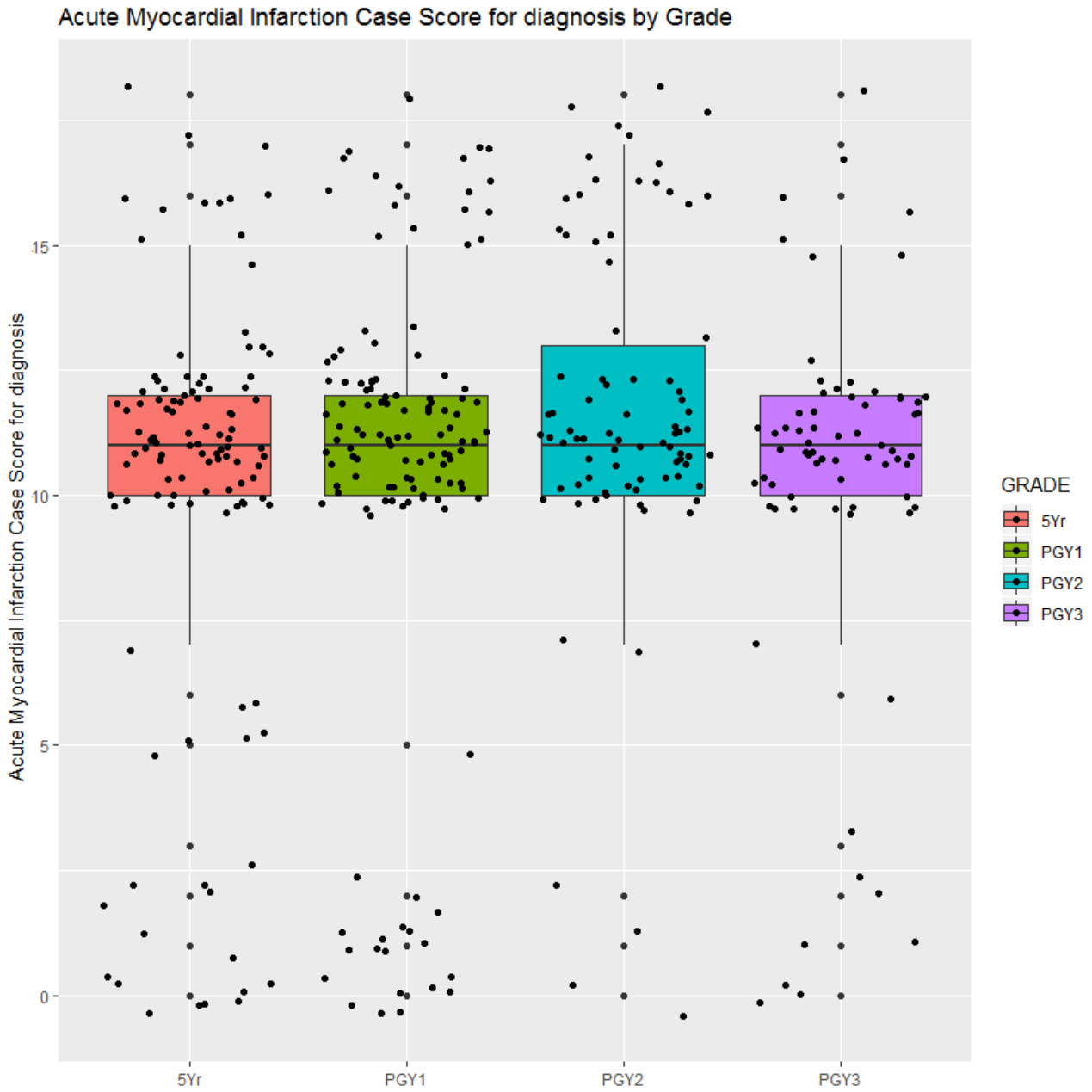
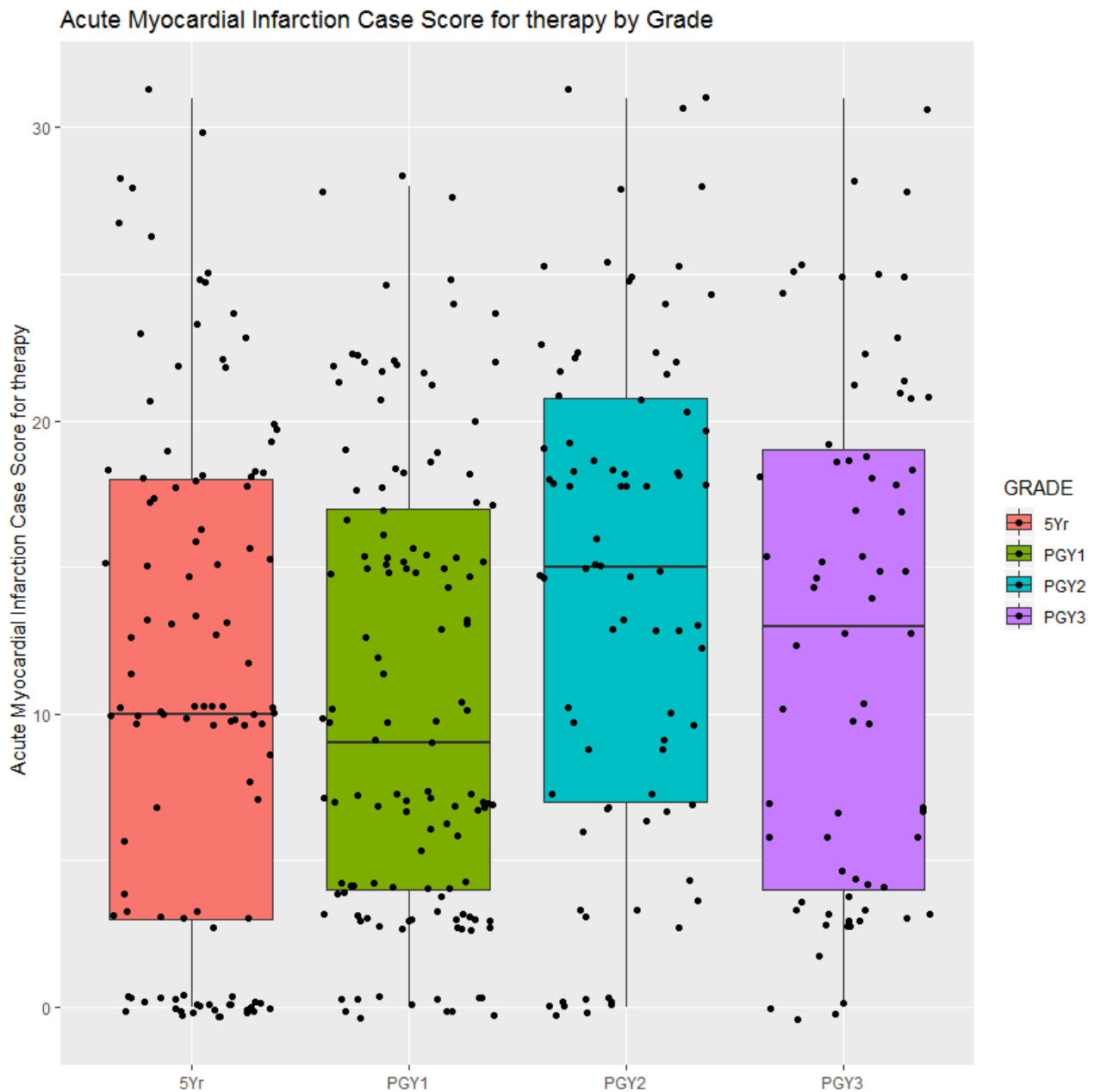


Figure 5. Subscore for diagnosis of students over training years.





**Figure 6.** Subscore for treatment of students over training years.

## Discussion

### Principal Findings

Before applying an assessment tool for use with medical students, we must obtain evidence for the instrument's reliability and validity [26-28]. Providing evidence of the validity of CBCRT will help the test management organization understand the effectiveness of the test from a broad and comprehensive perspective, clarify the aspects that the CBCRT can and cannot measure, and hence, allow for its continuous improvement. This is the goal of our current study. Our hypothesis was supported by the results obtained; specifically, senior students displayed higher testing scores than junior students (Table 3, Figure 1). In other words, the CMU-CBCRT is able to determine CR skills over different levels of medical education, especially in the early stage of the students' medical careers.

Looking specifically into the 4 categories of skills that we tested, we found that the most significant differences were revealed in the information collection, diagnosis, and treatment scores among junior and senior medical students. This was as predicted. With years of training, their experience and ability to clinically reason are improving, and as a result, they performed better on the information collection, diagnosis, and treatment, as well as the total CBCRT score. This further suggests that the CMU-CBCRT can determine the CR skills of students at varying levels.

We also carefully studied and analyzed why there were no significant differences in treatment error scores among the 4 training groups. For a simulated case of myocardial infarction, we can observe from the test result the challenge faced by participants who have never experienced this form of examination before. When the passing score was set at 60%,



the average score in this case (59.01 [SD 16.68]) did not pass. The choice of wrong treatment is a negative item in the scoring system, so the item writing expert is very cautious in formulating the scoring standard. Only behavior that caused extreme consequences resulted in points being deducted, and the weight was set at a very low level (ie, -1%). In this test, we observed that treatment error behavior happened more with junior students than senior students, although without statistical significance.

In the absence of an available gold standard for measuring CR, evidence for construct validity is sought after in this area of research. This is an ongoing process, in which the skill measured by the assessment tool is linked to some other attribute by a hypothesis or construct. With the development of validity theory, the validity concept has a new connotation and forms a method based on multilevel evidence [22]. Validity is no longer an attribute of the measurement tool itself but rather the extent to which the evidence collected supports the interpretation, inference, and decision making of the test score [27,29].

With the positive evidence presented, we should still be aware that validity verification is a dynamic process [27] and no education instrument is 100% effective [27]. Even if the evidence indicates that the validity of a course test is significant, the validity study must continue along with the development of

the CBCRT system. There are still many problems to be solved, such as the setting of the evidence framework for the specific test validity, determination of the validity criteria, feasibility of the evidence collection method, and quantification of evidence data. This will require in-depth discussion by future researchers. We aim for constant examination of these issues in the process of developing a reliable and valid CR training model. In the future, we would include more simulated cases with a wide range of case difficulties and distribute CMU-CBCRT to more students to increase sample size. We would then carefully collect data on student performance and feedback. We also plan to add graphics and animation to enhance the interface design.

### Limitations

However, there were some limitations in our study to its generalizability. First, the respondents of the research were from only 5 medical institutions in China. Second, the findings of our study were limited by the representativeness and scale of the study population.

### Conclusions

We provided evidence of construct validity of the CMU-CBCRT. It is able to determine CR skills over different levels of medical education, especially in the early stage of the students' medical careers.

---

### Acknowledgments

Authors thank the National Medical Examination Center of China, which sponsored this research through the funded project Sustainable Development of Computer-based Case Simulations Examination System. We thank educators and collaborators from China Medical University, Fudan University School of Medicine, Sun Yat-sen University School of Medicine, Xuzhou Medical University, and Binzhou Medical College who provided of testing places and expertise that greatly assisted the research. Special thanks to Ms Kritika Taparia for proofreading the manuscript and provide valuable comments.

---

### Authors' Contributions

TZ, XG, and BQ collected the data. TZ, BZ, and BS analyzed the data and wrote the first draft of the manuscript.

---

### Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

Scoring sheet and scoring rubrics.

[\[XLSX File \(Microsoft Excel File\), 11 KB-Multimedia Appendix 1\]](#)

---

### Multimedia Appendix 2

China Medical University-computer-based clinical reasoning testing data.

[\[XLSX File \(Microsoft Excel File\), 113 KB-Multimedia Appendix 2\]](#)

---

### References

1. Bendapudi NM, Berry LL, Frey KA, Parish JT, Rayburn WL. Patients' perspectives on ideal physician behaviors. *Mayo Clin Proc* 2006 Mar;81(3):338-344. [doi: [10.4065/81.3.338](#)] [Medline: [16529138](#)]
2. Ellaway R, Masters K. AMEE Guide 32: e-Learning in medical education Part 1: learning, teaching and assessment. *Med Teach* 2008 Jun 03;30(5):455-473. [doi: [10.1080/01421590802108331](#)] [Medline: [18576185](#)]
3. Schreuder HWR, Oei G, Maas M, Borleffs JCC, Schijven MP. Implementation of simulation in surgical practice: minimally invasive surgery has taken the lead: the Dutch experience. *Med Teach* 2011;33(2):105-115. [doi: [10.3109/0142159X.2011.550967](#)] [Medline: [21275542](#)]
4. Bliss M, Osler W. *A Life in Medicine*. Oxford: Oxford University Press; 1999.

5. Cook DA, Triola MM. Virtual patients: a critical literature review and proposed next steps. *Med Educ* 2009 Apr;43(4):303-311. [doi: [10.1111/j.1365-2923.2008.03286.x](https://doi.org/10.1111/j.1365-2923.2008.03286.x)] [Medline: [19335571](https://pubmed.ncbi.nlm.nih.gov/19335571/)]
6. Kassirer JP. Teaching clinical reasoning: case-based and coached. *Acad Med* 2010 Jul;85(7):1118-1124. [doi: [10.1097/acm.0b013e3181d5dd0d](https://doi.org/10.1097/acm.0b013e3181d5dd0d)] [Medline: [20603909](https://pubmed.ncbi.nlm.nih.gov/20603909/)]
7. Sevdalis N, McCulloch P. Teaching evidence-based decision-making. *Surg Clin North Am* 2006 Feb;86(1):59-70. [doi: [10.1016/j.suc.2005.10.008](https://doi.org/10.1016/j.suc.2005.10.008)] [Medline: [16442420](https://pubmed.ncbi.nlm.nih.gov/16442420/)]
8. Sokol RG, Slawson DC, Shaughnessy AF. Teaching evidence-based medicine application: transformative concepts of information mastery that foster evidence-informed decision-making. *BMJ Evid Based Med* 2019 Aug;24(4):149-154. [doi: [10.1136/bmjebm-2018-111142](https://doi.org/10.1136/bmjebm-2018-111142)] [Medline: [31208984](https://pubmed.ncbi.nlm.nih.gov/31208984/)]
9. Mottola MF, Bagnall KM, Belcastro AN. Effects of strenuous maternal exercise on fetal organ weights and skeletal muscle development in rats. *J Dev Physiol* 1989 Feb;11(2):111-115. [Medline: [2778291](https://pubmed.ncbi.nlm.nih.gov/2778291/)]
10. Peile E. Teaching balanced clinical decision-making in primary care: evidence-based and values-based approaches used in conjunction. *Educ Prim Care* 2014 Mar;25(2):67-70. [doi: [10.1080/14739879.2014.11494248](https://doi.org/10.1080/14739879.2014.11494248)] [Medline: [24593967](https://pubmed.ncbi.nlm.nih.gov/24593967/)]
11. Scott PJ, Altenburger PA, Kean J. A collaborative teaching strategy for enhancing learning of evidence-based clinical decision-making. *J Allied Health* 2011;40(3):120-127. [Medline: [21927777](https://pubmed.ncbi.nlm.nih.gov/21927777/)]
12. Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. *Med Educ* 2015 Oct 18;49(10):961-973. [doi: [10.1111/medu.12775](https://doi.org/10.1111/medu.12775)] [Medline: [26383068](https://pubmed.ncbi.nlm.nih.gov/26383068/)]
13. Furze J, Kenyon LK, Jensen GM. Connecting classroom, clinic, and context: clinical reasoning strategies for clinical instructors and academic faculty. *Pediatr Phys Ther* 2015;27(4):368-375. [doi: [10.1097/PEP.0000000000000185](https://doi.org/10.1097/PEP.0000000000000185)] [Medline: [26397080](https://pubmed.ncbi.nlm.nih.gov/26397080/)]
14. Windish DM. Teaching medical students clinical reasoning skills. *Acad Med* 2000 Jan;75(1):90. [doi: [10.1097/00001888-200001000-00022](https://doi.org/10.1097/00001888-200001000-00022)] [Medline: [10667884](https://pubmed.ncbi.nlm.nih.gov/10667884/)]
15. Custers EJFM. Theories of truth and teaching clinical reasoning and problem solving. *Adv Health Sci Educ Theory Pract* 2019 Oct;24(4):839-848 [FREE Full text] [doi: [10.1007/s10459-018-09871-4](https://doi.org/10.1007/s10459-018-09871-4)] [Medline: [30671703](https://pubmed.ncbi.nlm.nih.gov/30671703/)]
16. Levin M, Cennimo D, Chen S, Lamba S. Teaching clinical reasoning to medical students: a case-based illness script worksheet approach. *MedEdPORTAL* 2016 Aug 26;12:10445 [FREE Full text] [doi: [10.15766/mep.2374-8265.10445](https://doi.org/10.15766/mep.2374-8265.10445)] [Medline: [31008223](https://pubmed.ncbi.nlm.nih.gov/31008223/)]
17. van Merriënboer JJG, Sweller J. Cognitive load theory in health professional education: design principles and strategies. *Med Educ* 2010 Jan;44(1):85-93. [doi: [10.1111/j.1365-2923.2009.03498.x](https://doi.org/10.1111/j.1365-2923.2009.03498.x)] [Medline: [20078759](https://pubmed.ncbi.nlm.nih.gov/20078759/)]
18. Rencic J. Twelve tips for teaching expertise in clinical reasoning. *Med Teach* 2011;33(11):887-892. [doi: [10.3109/0142159X.2011.558142](https://doi.org/10.3109/0142159X.2011.558142)] [Medline: [21711217](https://pubmed.ncbi.nlm.nih.gov/21711217/)]
19. Forrest JL, Miller SA, Newman MG. Teaching evidence-based decision-making versus experience-based dentistry. *Alpha Omega* 2004 Jul;97(2):35-41. [Medline: [18712148](https://pubmed.ncbi.nlm.nih.gov/18712148/)]
20. Stone J, Moskowitz GB. Non-conscious bias in medical decision making: what can be done to reduce it? *Med Educ* 2011 Aug;45(8):768-776. [doi: [10.1111/j.1365-2923.2011.04026.x](https://doi.org/10.1111/j.1365-2923.2011.04026.x)] [Medline: [21752073](https://pubmed.ncbi.nlm.nih.gov/21752073/)]
21. Toro-Huamanchumo CJ, Arce-Villalobos LR. The clinical eye: a need to improve the teaching of semiology in undergraduate medical education. *Int J Med Students* 2014 Oct 25;2(3):144. [doi: [10.5195/ijms.2014.186](https://doi.org/10.5195/ijms.2014.186)]
22. Ward RC, Muckle TJ, Kremer MJ, Krogh MA. Computer-based case simulations for assessment in health care: a literature review of validity evidence. *Eval Health Prof* 2017 Jan 01:163278717718609. [doi: [10.1177/0163278717718609](https://doi.org/10.1177/0163278717718609)] [Medline: [28727944](https://pubmed.ncbi.nlm.nih.gov/28727944/)]
23. Allard M, Lafleur A, Richard E, Lebouthillier A, Vailles C. How medical students edited an OSCE Study Guide and why should you? *Int J Med Students* 2018 May 09;6(2):78-82. [doi: [10.5195/ijms.2018.248](https://doi.org/10.5195/ijms.2018.248)]
24. Nekkanti S, Manjunath S, Mahtani A, Meka A, Rao T. A survey based feedback analysis of the current medical teaching methodology and trends in medical research practice in a South Indian medical institute. *Int J Med Students* 2018 Apr 30;6(1):6-17. [doi: [10.5195/ijms.2018.3](https://doi.org/10.5195/ijms.2018.3)]
25. Liaw SY, Rashasegaran A, Wong LF, Deneen CC, Cooper S, Levett-Jones T, et al. Development and psychometric testing of a Clinical Reasoning Evaluation Simulation Tool (CREST) for assessing nursing students' abilities to recognize and respond to clinical deterioration. *Nurse Educ Today* 2018 Mar;62:74-79. [doi: [10.1016/j.nedt.2017.12.009](https://doi.org/10.1016/j.nedt.2017.12.009)] [Medline: [29306102](https://pubmed.ncbi.nlm.nih.gov/29306102/)]
26. Feldman LS, Sherman V, Fried GM. Using simulators to assess laparoscopic competence: ready for widespread use? *Surgery* 2004 Jan;135(1):28-42. [doi: [10.1016/s0039-6060\(03\)00155-7](https://doi.org/10.1016/s0039-6060(03)00155-7)]
27. Morris SB, Kwaske IH, Daisley RR. The validity of individual psychological assessments. *Ind Organ Psychol* 2015 Jan 07;4(3):322-326. [doi: [10.1111/j.1754-9434.2011.01347.x](https://doi.org/10.1111/j.1754-9434.2011.01347.x)]
28. Yu CHY, Straus S, Brydges R. The ABCs of DKA: development and validation of a computer-based simulator and scoring system. *J Gen Intern Med* 2015 Sep;30(9):1319-1332 [FREE Full text] [doi: [10.1007/s11606-015-3273-y](https://doi.org/10.1007/s11606-015-3273-y)] [Medline: [26173518](https://pubmed.ncbi.nlm.nih.gov/26173518/)]
29. Kane MT. Current concerns in validity theory. *J Educational Measurement* 2001 Dec;38(4):319-342. [doi: [10.1111/j.1745-3984.2001.tb01130.x](https://doi.org/10.1111/j.1745-3984.2001.tb01130.x)]

---

**Abbreviations**

**ANOVA:** analysis of variance  
**CBCRT:** computer-based clinical reasoning testing  
**CMU:** China Medical University  
**CR:** clinical reasoning  
**PBL:** problem-based learning  
**PGY:** postgraduate year

---

*Edited by N Zary; submitted 02.01.20; peer-reviewed by M Gaman, R Miclaus; comments to author 26.10.20; revised version received 19.03.21; accepted 19.06.21; published 09.11.21*

*Please cite as:*

Zuo T, Sun B, Guan X, Zheng B, Qu B

*Evidence of Construct Validity of Computer-Based Tests for Clinical Reasoning: Instrument Validation Study*

*JMIR Serious Games 2021;9(4):e17670*

URL: <https://games.jmir.org/2021/4/e17670>

doi: [10.2196/17670](https://doi.org/10.2196/17670)

PMID:

©Tianming Zuo, Baozhi Sun, Xu Guan, Bin Zheng, Bo Qu. Originally published in JMIR Serious Games (<https://games.jmir.org>), 09.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on <https://games.jmir.org>, as well as this copyright and license information must be included.