Students Who Have Taken a Practice Examination Achieve Better Computer-based Testing Results than Those Who Have Not

YUICHI TOMIKI^{*1) 2)}, TSUTOMU SUZUKI^{*1) 3)}, KAZUO KEMPE^{*1) 4)}, HIROO WADA^{*1) 5)},
 MAKINO WATANABE^{*1) 6)}, MIWA SEKINE^{*1) 7)}, YUKO TAKEDA^{*1)}, TAKAO OKADA^{*1)}

^{*1)}Division of Medical Education, Juntendo University Faculty of Medicine, Tokyo, Japan, ^{*2)}Department of Coloproctological Surgery, Juntendo University Faculty of Medicine, Tokyo, Japan, ^{*3)}Department of Respiratory Medicine, Juntendo University Faculty of Medicine, Tokyo, Japan, ^{*4)}Department of Internal Medicine and Rheumatology, Juntendo University Faculty of Medicine, Tokyo, Japan, ^{*5)}Department of Public Health, Juntendo University Faculty of Medicine, Tokyo, Japan, ^{*6)}Department of Organ and Cell Physiology, Juntendo University Faculty of Medicine, Tokyo, Japan, ^{*7)}Department of Microbiology, Juntendo University Faculty of Medicine, Tokyo, Japan

Objective: The present study examined the relationship between Computer-based Testing (CBT) results and the status of taking a practice examination.

Materials: A total of 392 (275 males and 117 females) students, who had taken CBT during their fourth year at Juntendo University within a 3-year period between 2017 and 2019.

Methods: We divided into 2 groups: those who had (practice examination group) and those who had not (non-practice examination group) taken a practice examination prior to CBT, and their Item Response Theory (IRT) scores were compared. IRT scores lower than 400 were defined as poor CBT results, and were compared between the 2 groups.

Results: The numbers of practice and non-practice examination group members were 228 (58.2%) and 164 (41.8%), respectively. The mean IRT score was significantly higher in the practice examination group (515.9 ± 88.6) than in the non-practice examination group (478.0 ± 78.3) ($p < 0.0001$). The proportion of students with IRT scores lower than 400 was significantly larger in the non-practice examination group ($p = 0.0004$).

Conclusions: Students who had taken a practice examination achieved better CBT results than those who had not. However, the status of taking a practice examination is only an index of preparedness for CBT, and does not indicate improved CBT results. Since CBT requires early preparation, delays in initiating test studies may result in insufficient studies and a superficial understanding.

Key words: common achievement test, computer-based testing (CBT), practice examination, Item Response Theory (IRT)

Introduction

Passing a common achievement test is a requirement for medical students to participate in clinical exercise. This test consists of a computer-based testing (CBT) as a knowledge evaluation and an Objective Structured Clinical Examination (OSCE) as a skill/attitude assessment¹⁾. CBT comprises 320 questions that are randomly presented from a pool

of questions. Questions presented vary among test takers; however, the level of difficulty is adjusted based on the Item Response Theory (IRT). The passing score is unified among all medical schools. Students with IRT scores lower than 359 do not pass the test, and those who fail to pass it again are not allowed to participate in clinical exercise²⁾. Some universities set the passing score at 359 or higher.

Corresponding author: Yuichi Tomiki (ORCID: orcid.org/0000-0001-9284-905X)

Division of Medical Education, Juntendo University Faculty of Medicine

2-1-1 Hongo, Bunkyo-ku, Tokyo 113-8421, Japan

TEL: +81-3813-3111 E-mail: tomiki@juntendo.ac.jp

[Received Nov. 28, 2019] [Accepted Jan. 23, 2020]

J-STAGE Advance published date: Feb. 18, 2020

Copyright © 2020 The Juntendo Medical Society. This is an open access article distributed under the terms of Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original source is properly credited.
 doi: 10.14789/jmj.2020.66.JMJ19-Re03

CBT examines basic knowledge of basic medicine and clinical medicine. Since these areas are extensively learned, preparation for the test needs to be initiated early. CBT is conducted immediately after the summer vacation at our university, and those who begin to study at the beginning of the summer vacation rarely pass.

Most students prepare for CBT using the workbooks published by a publisher specializing in medicine. When explanations in the workbook are insufficient for knowledge acquisition, students check their lecture notes or textbooks. Purchasers of the workbook, who have registered with an online membership system, are allowed to take a practice examination for free as a special benefit. An increasing number of students have recently been taking this practice examination prior to CBT to confirm their achievement levels. The present study examined the relationship between CBT results and the status of taking a practice examination.

Materials and Methods

A total of 392 (275 males and 117 females) students, who had taken CBT during their fourth year at Juntendo University within a 3-year period between 2017 and 2019, were divided into 2 groups: those who had (practice examination group) and those who had not (non-practice examination group) taken a practice examination prior to CBT. The students took or did not take the practice examination based on their own decisions. Therefore, the grouping was randomized.

The IRT score was compared between the practice and non-practice examination groups.

IRT scores for CBT are classified into 6 grades: 1: lower than 300, 2: 300–399, 3: 400–499, 4: 500–599, 5: 600–699, and 6: 700 or higher.

In the present study, IRT scores lower than 400 (Grades 1 and 2) were defined as poor CBT results for comparison between the 2 groups. IRT scores lower than 400 were defined as poor CBT results for comparison between the 2 groups.

All values are shown as the mean $\pm$ SD (range). Statistical analyses were performed using the t-test and chi-squared test in addition to JMP 9.0 (SAS Institute Inc., Cary, NC, USA). The significance level was set at $p < 0.05$ in all settings.

Students' test results were processed to prevent

the identification of individuals, and their written consent for the use of these data for the study was previously obtained.

The present study was conducted with the approval of the Ethics Committee of Juntendo University (Approval number: 2018129).

Results

1. IRT scores of practice and non-practice examination groups

The numbers of practice and non-practice examination group members were 228 (58.2%) and 164 (41.8%), respectively. On analysis using a box plot, the interquartile range for the practice examination group was higher than that for the non-practice examination group. The scores of some students belonging to the practice examination group were so high that they were counted as outliers. In contrast, many students belonging to the non-practice examination groups achieved scores lower than 400, and their minimum value did not reach 300. The mean IRT score was significantly higher in the practice examination group (515.9 ± 88.6) than in the non-practice examination group (478.0 ± 78.3) ($p < 0.0001$) (Figure-1).

2. Status of taking a practice examination among students with IRT scores lower than 400

Thirty-six (9.2%) students had IRT scores lower than 400. Among these students, 11 (4.8%) and 25 (15.2%) belonged to the practice and non-practice examination groups, respectively. The proportion of students with IRT scores lower than 400 was significantly larger in the non-practice examination group ($p = 0.0004$) (Table-1).

Discussion

Based on the model core curriculum, CBT examines basic knowledge of basic medicine, clinical medicine, and social medicine, and serves as a mid-term examination during the 6-year medical school period³⁾. Additionally, studying to prepare for CBT before the initiation of clinical exercise is a good opportunity for students to organize the knowledge they have acquired through previous learning. Since this preparation needs to be conducted simultaneously with studies for the regular curriculum and regular tests, daily efforts are important. Students are also required to show

Figure-1 IRT scores of Practice and Non-practice Examination Groups

Table-1 Status of Taking a Practice Examination among Students with IRT Scores Lower than 400

IRT score <400		
Practice examination group	(n=228)	11 (4.8%)
Non-practice examination group	(n=164)	25 (15.2%)

p=0.0004

appropriate attitudes toward the test. CBT is conducted immediately after the summer vacation at our university, and those who begin to study at the beginning of the summer vacation rarely pass. Thus, the period when test studies are initiated and the duration of preparation are important factors affecting CBT results. However, some students lack the sense of importance about their academic records, and, thus, do not initiate preparation for the test despite its importance for their advancement. It is clear that delays in initiating test studies shorten the duration of preparation.

Absolutely insufficient studies and a superficial understanding lead to failures in CBT. Absolutely insufficient studies result from delays in their

initiation. When running out of time for preparation, students begin to simply resolve practice questions, rather than studying, which results in superficial knowledge without a deep understanding as a vicious circle. On the other hand, some students lack a sufficient understanding despite a sufficient duration of preparation because they repeatedly resolve practice questions. When explanations in the workbook are insufficient for understanding, students need to check their lecture notes or textbooks or ask faculty members or friends to clarify their unclear points. A superficial understanding also prevents the use of knowledge in clinical exercise⁴⁾.

The present study examined students' CBT

results only by comparing those taking and not taking a practice examination. Among all CBT takers, 58.2% had taken a practice examination prior to the test. Our university does not force students to take a practice examination. Students themselves decided whether to take the practice examination, and those who decided not to take it did not necessarily lack seriousness toward their studies. However, the number of these students is markedly increasing.

In the present study, CBT results were better in the practice examination group than in the non-practice examination group, which indicates that many of the former had more time to spare during preparation; therefore, they took a practice examination.

By taking a practice examination, they may have become able to objectively assess their achievement levels, address their weak points, and consequently achieve better test results. In contrast, in the non-practice examination group, “I could not access the online membership system because my workbook had been handed down from a senior student” or “I could access the online membership system, but had no courage to confirm my achievement level” may have been the main reason for not taking a practice examination. In the case of many students with poor CBT results, a lack of time may have prevented them from taking the practice examination. It is important to note that the status of taking a practice examination is only an index of preparedness for CBT, and does not indicate improved CBT results. At the same time, taking a practice examination is effective for test takers to habituate themselves to the style of the CBT. For example, in the last part of CBT, questions are presented in a specific format. Four questions are consecutively presented for each case, and test takers need to resolve them in order, but cannot go back to the previous question after confirming a solution and advancing to the next question¹⁾. Thus, habituation to the test style may explain the better results obtained by the practice examination group.

As a study limitation, the results of analyses were

obtained from a single facility, and data regarding the practice examination were also provided by a single publisher. To confirm the results of the present study, further studies should be conducted, involving multiple facilities and publishers publishing workbooks for CBT. In addition, CBT results were poor among some practice examination group members and favorable among some non-practice examination group members, indicating the necessity of examining other factors associated with poor CBT results. Changes in CBT results with an increase in the number of students taking the practice examination are also an interesting topic for future studies.

When minimum passing scores were collected from all medical schools in 2015, students and universities both began to feel more pressure about CBT. Since CBT requires early preparation, delays in initiating test studies may result in insufficient studies and a superficial understanding. In the future, the causes of delays in initiating test studies also need to be examined as a basis for more effective education and guidance for students.

Conflict of interest

No conflicts of interest were declared by any of the authors.

Acknowledgments

The authors would like to express their gratitude to Medic Media Co., Ltd. for offering practice examination data.

References

- 1) Common Achievement Tests Organization: Common achievement tests before clinical exercise, 16th edition. 2018. (<http://www.cato.umin.jp/>)
- 2) Nakaya H: Medical Education in Japan and National Medical Practitioner Examination. Medical Education, 2015; 46: 14-17.
- 3) Ministry of Education, Culture, Sports, Science and Technology/Japan: Model core curriculum of medical education. (<https://www.mext.go.jp/>)
- 4) Tomiki Y, Dambara T, Okada T, *et al*: Do academic rankings based on common achievement tests predict the results of graduation examination results? Juntendo Medical Journal, 2012; 58: 431-435.