

Item analyses of 2016 written exam of Turkish Society of Cardiovascular Surgery Proficiency Board

*Türk Kalp ve Damar Cerrahisi Derneği Yeterlik Kurulu
2016 yazılı sınavı madde analizleri*

Hatice Şahin¹, Yüksel Atay², Tahir Yağdı², Serap Aykut Aka³

Institution where the research was done:
Medicine Faculty of Ege University, İzmir, Turkey

Author Affiliations:

¹Department of Medical Education Education, Medicine Faculty of Ege University, İzmir, Turkey

²Department of Cardiovascular Surgery, Medicine Faculty of Ege University, İzmir, Turkey

³Department of Cardiovascular Surgery, Dr. Siyami Ersek Thoracic and Cardiovascular Surgery Training and Research Hospital, İstanbul, Turkey

ABSTRACT

Background: This study aims to evaluate the item analyses of the 2016 written exam of Turkish Society of Cardiovascular Surgery Proficiency Board and conduct yearly comparisons.

Methods: The last of the written exams, which are performed annually, were held in Eskişehir in May 2016. The exam items were analyzed in terms of difficulty, discrimination, and validity.

Results: Correct answer items were distributed equally. The average percentage of questions answered correctly was 70.7%. Of the questions, 56% were answered above the average. Range of exam scores was narrower compared to last year's exam. According to the item analysis; rate of questions with the desired difficulty level was 39%, rate of questions with the desired discrimination level was 52%, rate of invalid questions was 54%, and rate of questions with high level of validity was 9%. A comparison of the exams of years 2015 and 2016 revealed that the latter exam was easier, had low discrimination, and weaker validity.

Conclusion: According to our findings, subject areas and number of questions in the exam set were effective on the validity of the exam. At least 70% of the questions in the exam set should include the main topics of the specialty and the number of questions should not be changed according to years. With the Proficiency Board forming a question bank and keeping the questions with high validity in the bank, the exam may become more qualified over the years.

Keywords: Board exam; cardiovascular surgery; item analysis.

ÖZ

Amaç: Bu çalışmada Türk Kalp ve Damar Cerrahisi Derneği Yeterlik Kurulunun 2016 yazılı sınavının madde analizleri değerlendirildi ve yıllar arası karşılaştırmalar yapıldı.

Çalışma planı: Yıllık olarak yapılan yazılı sınavların sonucusu Mayıs 2016'da Eskişehir'de gerçekleştirildi. Sınav maddeleri zorluk, ayırıcılık ve geçerlilik açısından analiz edildi.

Bulgular: Doğru yanıt şıkları eşit dağıtılmıştı. Soruları doğru yanıtlama oranı ortalama %70.7 idi. Soruların %56'sı ortalamanın üzerinde yanıtlanmıştı. Sınavdan alınan puan aralığı geçen yıl yapılan sınava göre daha dar idi. Madde analizine göre; istenen zorluk düzeyine sahip soru oranı %39, istenen ayırıcılık düzeyine sahip soru oranı %52, geçersiz soru oranı %54 ve yüksek düzeyde geçerli soru oranı %9 idi. 2015 ve 2016 yılları sınavlarının karşılaştırılması son sınavın daha kolay, ayırıcılığının düşük ve geçerliliğinin daha zayıf olduğunu ortaya koydu.

Sonuç: Bulgularımıza göre sınav setindeki konu alanları ve soru sayıları sınavın geçerliği üzerinde etkili idi. Sınav setinde soruların en az %70'i uzmanlığın ana başlıklarını içermeli ve soru sayısı yıllara göre değiştirilmemelidir. Yeterlik Kurulunun bir soru bankası oluşturması ve geçerliliği yüksek soruları bankada koruması sayesinde sınavlar yıllar içinde sınavlar daha nitelikli hale gelebilir.

Anahtar sözcükler: Yeterlik sınavı; kalp ve damar cerrahisi; madde analizi.

Received: February 25, 2017 *Accepted:* April 24, 2017

Correspondence: Hatice Şahin, MD. Ege Üniversitesi Tıp Fakültesi, Tıp Eğitimi Anabilim Dalı, 35100 Bornova, İzmir, Turkey.

Tel: +90 232 - 390 18 44 e-mail: hatice.sahin@ege.edu.tr

Cite this article as:

Şahin H, Atay Y, Yağdı T, Aykut Aka S. Item analyses of 2016 written exam of Turkish Society of Cardiovascular Surgery Proficiency Board. Turk Gogus Kalp Dama 2017;25(4):600-7.

©2017 All right reserved by the Turkish Society of Cardiovascular Surgery.

Specialization training provided by different institutions and under the guidance of different trainers may cause differences in proficiencies of physicians even if the trainers have the same titles. Therefore, this situation causes troubles also by differentiating the presentation of healthcare services. Notwithstanding that learning via Apprenticeship Model is still have a valuable place, getting minimum proficiencies required by the section of specialization became one of the main areas of interest for specialization training institutions and healthcare authorities in recent years.^[1] Committee for Specialty in Medicine Curriculum Development and Standard Setting System (TUKMOS) conducts various studies to form curriculum and set standards for each specialty section.^[2] The purpose of the TUKMOS study is to establish a curriculum that will ensure that graduation competencies are achieved during specialist education and set the standards of educational institutions. Specialization societies in Medicine are also working to improve the quality of specialization training via continuous professional development activities (schools, courses etc.) and board examinations towards the section of specialization. Specialization societies evaluate whether the specialist physician and specialist physician candidate has taken the training related with the field of interest by the workings of board committees. Getting the board certification is assessed as an indicator that the physician has completed a standard training program and possesses knowledge, skills and experience related with the section of the specialty.^[1] Thereby, specialization societies constituted board committees and started to give board certification and monitoring as standardization, auditing and monitoring of training for specialty in medicine is increasingly gaining importance currently.^[1]

Functions of board committees are described in legal texts.^[3] Cardiovascular Surgery Board Committee has done 13 written, 8 verbal/administration examinations since it was founded. The number of people who received board certification is 393.

In this study, we aimed to give information regarding written examination analyses by Cardiovascular Surgery Board Committee in 2016 and to provide guidance for the committee for comparing the years and setting a board examination set.

PATIENTS AND METHODS

Last written examination was performed at 8th May 2016 in Eskisehir by the Cardiovascular Surgery Board Committee. Examination questions prepared by faculty

members who are specialists in their field and invited for preparing questions by Specialization Society were included into examination set and an examination, which allows the answers obtained via optic reader.

Cardiovascular surgery specialists and residents participated in 2016 written examination as required by board committee directive and 92 people submitted.

Preparation and application of the examination set

The written examination set was prepared to include key issues related with cardiovascular surgery (Table 1).

Individuals who are specialists in their field asked to prepare the examination questions as multiple-choice and single correct answer. Negative score calculation (wrong answer removing the correct one) is not done. Examination set as a single booklet has 100 questions with one score for each of them.

Score of each examinee in examination (absolute score) were calculated. Board committee determined getting minimum 65 score in examination as success criterion through reconciliation.^[4] Examinees were informed whether they succeeded in the examination instead of the score they got.

Item analysis

Item analysis is analysis of each item in scope of the examination by means of statistical technics after the examination was applied to a group.^[5] Item analysis; helps to determine defective items and to obtain information about the parts of the examination which requires to be improved.

Using item analysis, information regarding difficulty and distinctiveness level. Difficulty index is a measure of the proportion of examinees who

Table 1. Distribution of examination questions according to subject areas

Subject area	Number of questions
Vascular surgery	25
Valve surgery	21
Congenital heart surgery	15
Aortic surgery	13
Coronary surgery	7
Cardiopulmonary bypass techniques	6
Intensive care principles	6
Heart failure surgery	3
Cardiac anatomy	3
Arrhythmia surgery	1
<i>Total</i>	100

answered the item correctly. Discrimination index is a measure of how well an item is able to distinguish between examinees who are knowledgeable and those who are not. Two different technics used for calculation of the item analysis. In the first technic, index computed from equal-sized high and low scoring groups on the exam with using Truman Kelley's "27% of sample" group size. In this technic, difficulty and discrimination index calculated on questions. Item difficulty index may range between 0 and 1. The mean of difficulty indexes for entire exam are required to be around 0.5. Item discrimination index may range between -1 and +1. Discrimination index is required to be higher than 0.3.^[6] In this analysis, the data of 46% of the students in the group are not taken into consideration. The comment about the reliability of the exam set is incomplete. Therefore, in the second analysis named biserial analysis, difficulty index and validity index calculation performed upon all examinees taking the examination. Reliability gives information about applying and scoring the questions consistently, whereas the validity gives information about whether the exam actually measures the situation that it wants to measure. With both analyzes it is possible to obtain information about the validity and reliability of the test.

Statistical analysis

In this section, the distribution of scores (mean, median, standard deviation, interval, minimum and maximum value), corrects responding rate, average correct response rate of the questions, success distribution of examinee, distribution of correctly answered questions above and below success limit, distribution of correctly answered questions above and below success limit according to subject area, difficulty and discrimination index are calculated.

With these analyzes, validity and reliability indices were calculated according to the subject area of the exam set. In addition, descriptive statistics related with the examinees were also calculated. Chi-square was used when comparing success score with categorical variables and Student t-test when comparing with continuous variables. Alpha value was accepted as 0.05 for the evaluation of statistical significance. Data were calculated in PASW 18.0 version (SPSS Inc., Chicago, IL, USA) data package.

RESULTS

A total of 69.56% (n=64) of the examinees of 2016 board written examination prepared by Turkish Cardiovascular Surgery Board Committee took the examination. 89.1% (n=57) of the examinees were male and 10.9% (n=7) were female. Gender distribution was as it was expected when specialty training and occupational experience of cardiovascular surgery discipline is considered. 76.6% (n=49) of the examination examinee were specialist physician and 24.4% (n=15) were residents.

When evaluating the institutions that examination examinees are currently working; it was determined that 50% of them were working at universities, 31.3% were working at education and research hospitals and 9.4% of them were working at state hospitals. 6.3% of the examinees were working at private hospitals. 3% of the examinees did not mention the institution they work. In this case, 81.3% of the examinee were working at an institution that provides specialty training.

When the score distribution were reviewed, the mean was determined as 70.7±7.8 (min: 52, max: 83) scores. The median value, which was the peak value of the scores of the examinees, was 72. Range score is 31

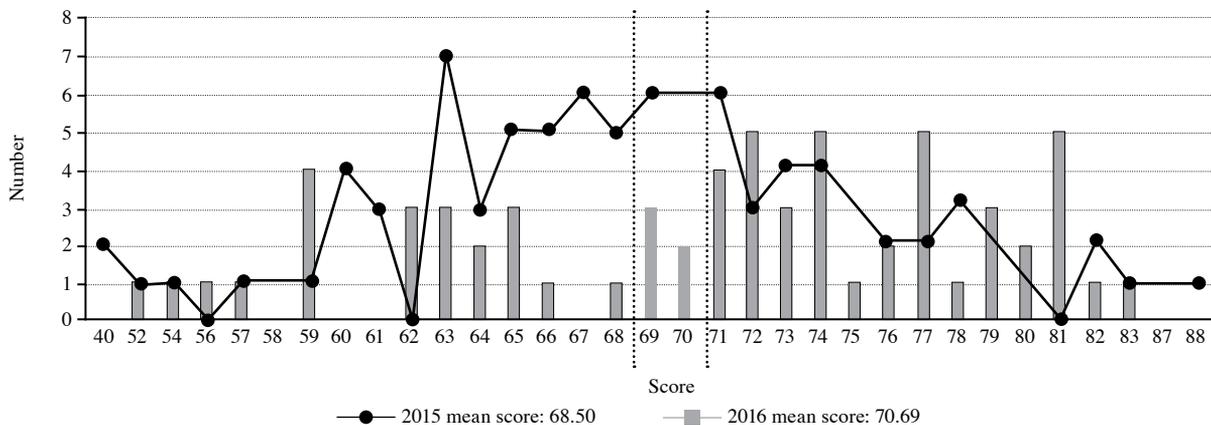


Figure 1. Comparison of examination scores according to years.

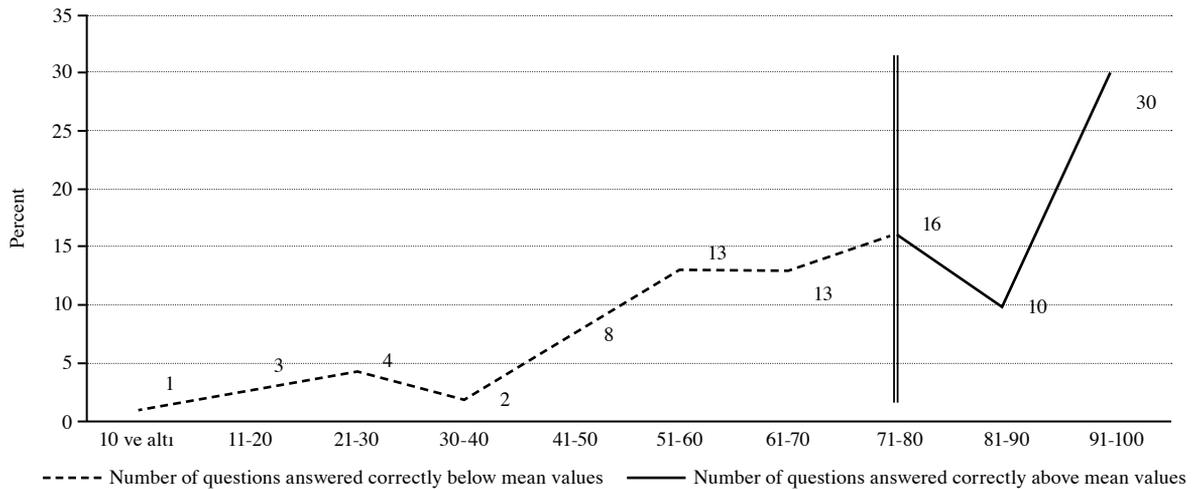


Figure 2. Distribution of right response rates of examination questions.

and it is less than half of the maximum score received. This is evidence that the scores on this year's exam distributed in a narrower range than the 2015 scores. For this reason, the distribution graph of the scores is right-aligned (Figure 1).

The success rate of the examination (65 scores and over) was 75% (n=48). No statistical significance was determined when the success rate according to the gender of the examinees (Chi-square=4.31 p=0.59). However, success rate of the males (78.9%) was higher than the success rate of female (42.9%).

No statistical significance was determined when examination success rate of the examinees according to the institutions they are working currently (Chi-square 1.05 p=0.90). 78.1% of the examinees who are working at universities and 75% of the examinees who are working at education and research hospitals were successful.

No statistical significance was determined when examination success rate of the examinees considering they are specialist or residents (Chi-square=0.02 p=0.86). Success rates of the specialists (75.5%) and residents (73.3%) were similar.

It was determined that the correct answers option were evenly distributed in the exam set (18-23%). This is an indicator for a meticulous approach in distribution of the right options in the examination set. In this way, the chance of the examinees to score by chance has been reduced. The mean correct response rate of examination questions was found as 70.7±23.2% (min: 5, max: 100). Forty four questions in the examination set (44%) were answered in a correct way below the

mean values. A total of 56 questions (56%) were answered in a correct way above the mean values (Figure 2).

Correct response rates of subject areas are calculated. Congenital heart surgery, valve surgery and vascular surgeries are highly responsive to questions. (Figure 3).

Findings related with item analysis

Difficulty index of the examination questions were calculated by comparison of 27% lower and upper groups. The mean difficulty index of questions was found as 0.7±0.2 (min: 0.05 max: 1.00). 39% of

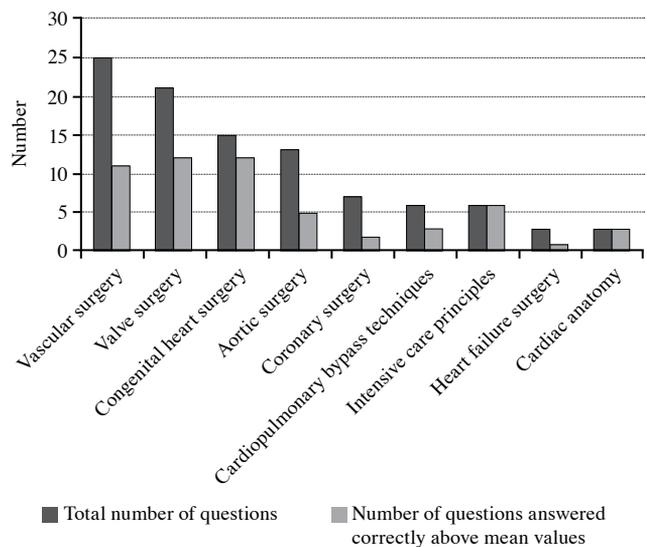


Figure 3. Distribution of answered questions correctly above the mean values according to subject area.

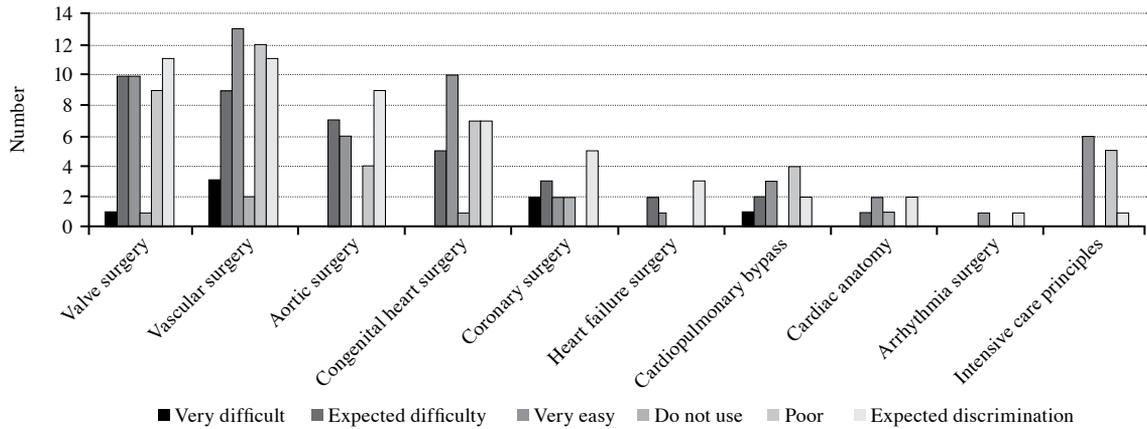


Figure 4. Difficulty and discrimination level of the questions according to subject areas.

the exam questions (39 questions) have the desired (easy, recommended, difficult) difficulty index. Fifty percent of the questions were very easy and 7% of them were very hard. The questions of valve surgery (10 questions), vascular surgery (9 questions) and aortic surgery (7 questions) were included in this scope. Arithmetic surgeon and intensive care unit subjects did not have a “difficulty index at the desired level”. All questions in these subjects were analyzed as easy questions.

Discrimination levels of the examination were calculated by comparison of 27% lower and upper group. The mean discrimination index of questions was found as 0.9 ± 0.2 (min: -0.24 max: 0.70). In the 52% of the questions (52 questions), the discrimination index was at the desired level (at the border, good and very good). Discrimination index was found weak at 41% of the questions. Seven percent of the questions

were analyzed as not to be used. Fifty two questions at the desired level include questions of valve surgery (11 questions), vascular surgery (11 questions), aortic surgery (9 questions) and congenital heart surgery (7 questions). A total of seven questions (valve surgery, congenital heart surgery, vascular surgery, coronary surgery and cardiac anatomy questions) were identified as “should not be used” in the analyzes. These questions had problems in ability of discriminating who knows and who does not know (Figure 4).

The validity of the exam set was examined by Biserial analysis. Nine percent of questions was found as highly validated, 19% of them mid-level validated, 17% of them partly validated and 54% of them invalid and 1% of them absolutely invalid. When validation status of the questions were compared according to their subject areas; vascular

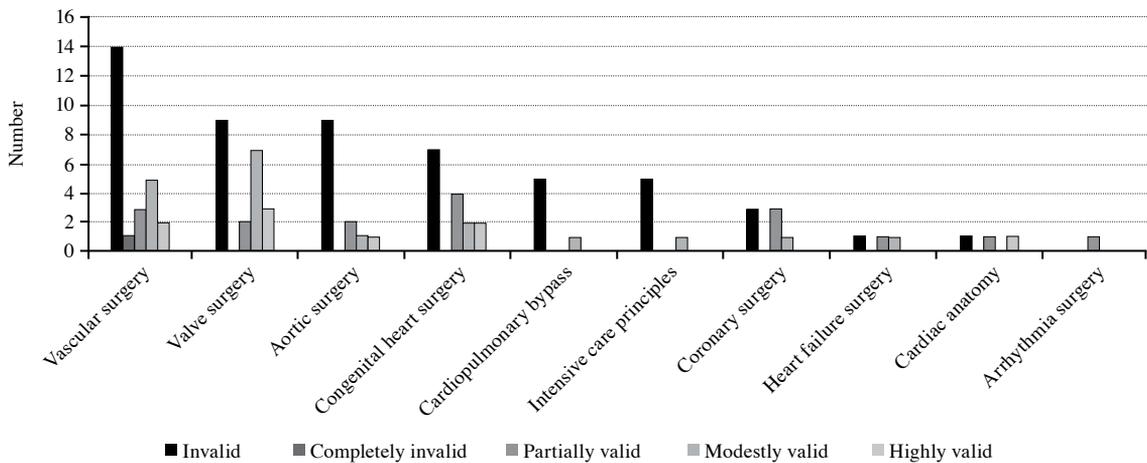


Figure 5. Validity level of the questions according to subject areas.

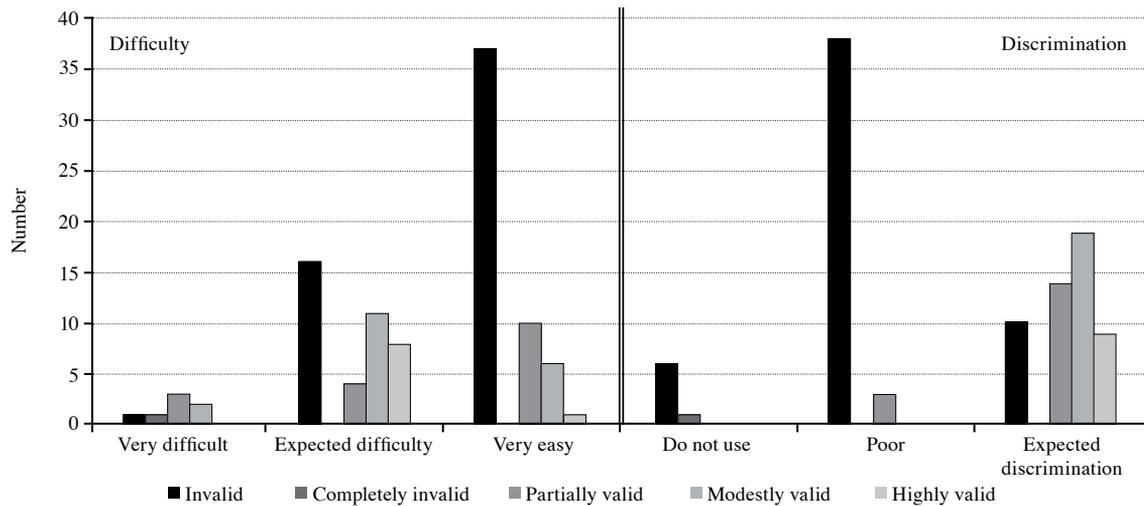


Figure 6. Difficulty and discrimination status of invalid questions.

surgery (14 questions), valve surgery (9 questions), aortic surgery (9 questions) and congenital heart surgery (7 questions) questions were found invalid. It was determined that these questions were not able to measure the subject area of information desired to measure (Figure 5).

One question in vascular surgery subject area was analyzed as absolutely invalid. When the reason for this was investigated, it was determined that the question was very difficult and answered very poorly, so the question reduced the validity index.

Subsequent analyzes were based on 54 questions that were analyzed as invalid done.

When the difficulties and discriminatory conditions of the invalid questions were evaluated, 37 (68.5%) of the invalid 54 questions were found to be very easy and 38 (70.3%) were poorly discriminated (Figure 6). Six (11.1%) of invalid questions were determined as “not to be used in analyses”. This was an indicator that analyses support each other.

The fact that there are invalid questions in the exam set explains the reason why the score graph is leaning to the right. Due to the fact that nearly half

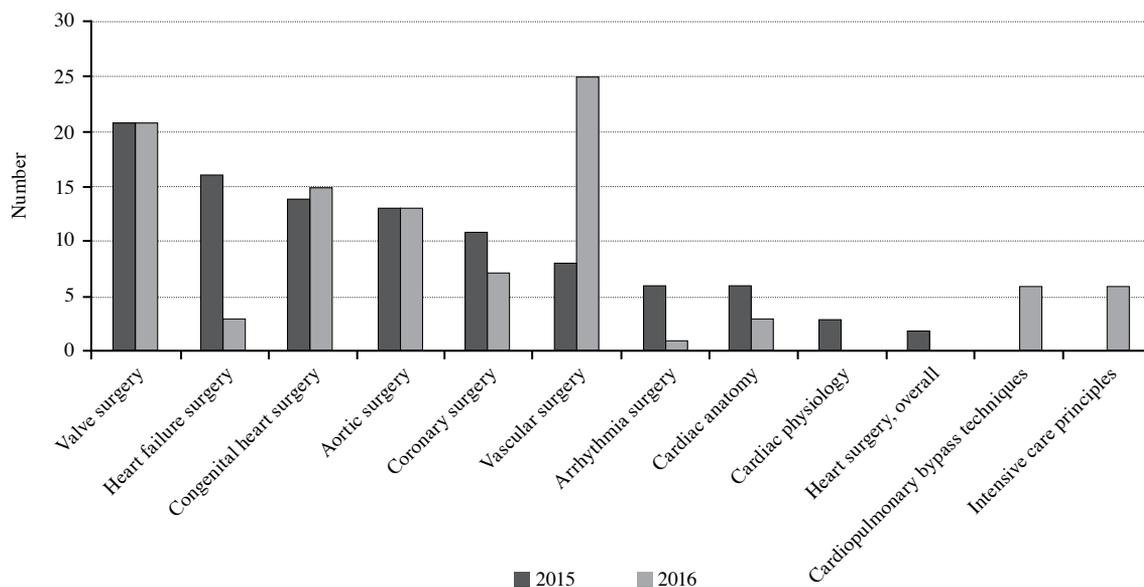


Figure 7. Comparison of the subject areas according to years.

of the questions on the exam set are very easy, these questions have been answered correctly by the majority of the examinees. For this reason, the power of discriminating between knowing and non-knowing has weakened. In sum, as the number of invalid questions increases, the correct answer rate of the examinees has increased and the scores from the exam have also increased. This also caused the examinee's score to cluster to the right of the average score.

Technical analysis of invalid questions was re-examined with a subject matter expert. During the preparation of the exam set, difficulties were encountered in obtaining subject-specific questions from subject matter experts. In order to overcome these difficulties, the questions used in the previous exams were included in the exam set with the same or slight changes. Due to these problems, the exam set is composed of more information-oriented and easier questions and the validity of the exam is negatively affected.

DISCUSSION

The first stage in the evaluation of the proficiency of the Turkish Society of Cardiovascular Surgery Board Committee is a written examination. The last written exam was held in Eskişehir on May 2016. Compared to the written exam analyzes made in 2015 and 2016, both sets included only one correct answer 100 questions.^[7] Subject areas were different in the two exams (Figure 7). While the number of valve surgery and congenital surgery questions was the same, the vascular surgery questions were numerous in the 2016 exam set. Questions about cardiopulmonary bypass techniques and intensive care principles are only available on the 2016 exam set. In the 2016 exam set, the weight of the exam questions according to subject areas was determined in determining the number of questions written in 2015.

In addition to these, determinations regarding written exams in 2015 and 2016 are summarized below;^[7]

- There has been less participation for the 2016 examination than in the previous year, but the participation of residents is higher in 2016.
- Physicians working in research hospitals participated more in the 2016 examination than in the 2015 examination.
- The average score of the 2016 exam is higher than the previous year.
- The desired level of difficulty according to subject areas has changed with the years.

- The desired level of difficulty according to subject areas has changed with the years. It was determined that the 2016 examination was made of very easy questions and the validity of the examination dropped.

The 2015 and 2016 Board written exam analyzes gave different levels of information about the reliability and validity of exams. Based on the findings of the analysis, the following proposals have been developed for the following examinations.

- For the exam set, clarification of subject areas and number of questions should be fixed as they are the same for each exam. A preliminary study has been done for this.
- For the next 100 question sets, 70 questions will be taken from the main subject areas [heart failure surgery (15 questions), valve surgery (15 questions), coronary surgery (15 questions), aortic surgery (15 questions) and congenital heart surgery (10 questions)] and 30 questions will be taken from the side subject areas [vascular surgery (10 questions), cardiopulmonary bypass techniques (7 questions), intensive care (6 questions) and cardiac anatomy (7 questions)]. In addition, the number of questions related to the main topics will be fixed during the examinations to be held in the following years. Thus, the exam coverage will be standard and the comparison over the years will be better.
- For the exams, it is necessary for the trainers to make an effort to prepare based on synthesis and assessment as well as preparation of knowledge-based questions.
- Board committee should examine exam analysis in detail. Particularly invalid questions should be reviewed and corrected. It should form a question bank for written exams.

Acknowledgment

We appreciate below mentioned members of board committee for their efforts in preparation and development of the examination.

- Suat Buket, Kent Hospital, Cardiovascular Surgery Clinic, İzmir
- Saadettin Dernek, Osmangazi University Faculty of Medicine Department of Thoracic and Cardiovascular Surgery, Eskişehir
- Murat Özeren, Mersin University Faculty of Medicine, Research and Application Hospital, Department of Cardiovascular Surgery, Mersin

- Tufan Paker, American Hospital, Department of Cardiovascular Surgery, İstanbul
- Sadık Eryılmaz, Ankara University Faculty of Medicine Cardiac Center Cardiovascular Surgery Clinic, Ankara
- Mehmet Kaplan, Dr. Siyami Ersek Thoracic and Cardiovascular Surgery Education and Research Hospital, Cardiovascular Surgery Clinic, İstanbul

Declaration of conflicting interests

The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

The authors received no financial support for the research and/or authorship of this article.

REFERENCES

1. Kınay M, Zorlu P. UYEK faaliyetleri ve TTB-UDKK-UYEK çalışmayı. TTB-UDKK Bülteni 2006;1;21-2.
2. Tıpta Uzmanlık Kurulu Müfredat Oluşturma ve Standart Belirleme Sistemi, Kalp ve Damar Cerrahisi Uzmanlık Eğitimi Çekirdek Müfredatı. Available from: http://www.tuk.saglik.gov.tr/muf2.1/kalp_ve_damar_cerrahisi/kalp_damar_cerrahisi_mufredat_v.2.1.pdf. [Accessed: February 22, 2017]
3. Türk Kalp ve Damar Cerrahisi Yeterlik Kurulu Yönergesi. Available from: <http://www.tkdcd.org/content/157/yeterlik-kurulu-yonergesi>. [Accessed: February 22, 2017].
4. McAleer S. Choosing assessment instruments. In: Dent John A, editor. A practical guide for medical teachers. Dundee: Elsevier; 2000. p. 279-88.
5. Tekin H, editör. Eğitimde ölçme ve değerlendirme. Ankara: Yargı Yayınları; 2000.
6. Ölçme araçlarında bulunması gereken nitelikler. Available at: <http://talimterbiye.mebnet.net/program-gel-birimi/Soru%20Analizleri-%20Rapor.pdf>. [Accessed: February 22, 2017].
7. Şahin H, Atay Y, Yağdı T, Aykut Aka S. Türk Kalp ve Damar Cerrahisi Derneği Yeterlik Kurulu 2015 yazılı sınavı madde analizleri. Turk Gogus Kalp Dama 2015;23:427-33.