Indian Journal of Surgery (April 2022) 84 (Suppl 1):S269–S274
https://doi.org/10.1007/s12262-022-03408-1

ORIGINAL ARTICLE

The Israeli Surgical Oral Board Examination: Improved Objectivity Increases Knowledge Authenticity

Guy Pines[1,2] · Ofir Ben-Ishay[3] · David Hochstein[3] · Guy Lahat[4] · Gil Ohana[5] · Gilbert Sebbag[6] · Yoram Kluger[3] · Ariel Halvey[7]

Received: 1 March 2022 / Accepted: 4 April 2022 /Published online: 14 April 2022

Abstract

The oral board examination in general surgery in Israel was recently revised aiming for improved objectivity and standardization. Herein, we describe the current exam model. Between 05/2018 and 11/2020, two exams per year were conducted with the current exam model. The examination consists of 12 stations, each focusing on a different field. Passing score is 80%. At the end of the examination, examiners and candidates complete a questionnaire regarding the examination's process and quality (scale 1–5). A total of 142 residents attended six oral board examinations. Mean pass rate was 76.6 ± 9.5%. Questions with overall highest pass rates were acute-care surgery (86.6 ± 4.8%), foregut surgery (84.6 ± 7.6%), and colorectal surgery (84 ± 8.1%). Questions with the highest fail rates were surgical oncology (31.7 ± 13.3%) and abdominal-wall surgery (28.8 ± 16.9%). Examiners' feedback scored highest the following: standardization of the exam (4.45 ± 0.63) and whether the presented cases reflect the daily work of an attending surgeon (4.35 ± 0.87). Candidates' feedback scored highest the following: did the examiners treat you in an appropriate manner (4.08 ± 1.17). In conclusion, oral exams are challenging and bear limitations, but properly constructed exams allow good evaluation of the trainees' thinking process and decision-making skills, without compromising exam's integrity and standardization.

Keywords Oral board examination · Trainee evaluation · Standardization · Objectivity

Introduction

In Israel, 34 general surgery wards offer general surgery residency programs. Only graduates of an accredited medical school with a valid license to practice medicine in Israel are eligible for these residency programs. The residency program in general surgery is approved by the Israeli Medical Association (IMA) and follows a strict curriculum. The general surgery residency program lasts 6 years, of which 4.5 are devoted to general surgery. Six months of research (basic science) and 3 months of intensive care are mandatory rotations. Three additional rotations of 3 months each are optional in any surgical sub-specialty.

Board certification requires completing 600 surgical procedures, passing a written board examination (obtainable after completing three years of residency) and passing an oral board examination (presented to residents in their sixth year of residency), along with the recommendation of the department head attesting to the resident's competency to practice as a senior and independent general surgeon. Recently, a major revision has been made to the mandatory oral board examination process. Herein, we describe the changes and evolution of the exams, and report candidates' and examiners' feedback along with success rates.

Historical Notes

The Israeli Medical Association published certification guidelines in 1976 for all medical specialties, which specified that residents must successfully pass both a written

Guy Pines, Yoram Kluger and Ariel Halvey contributed equally to this work.

This study was presented in the Israeli medical education conference, Nov 2021.

✉ Guy Pines
guypi@clalit.org.il; guypines@hotmail.com

✉ Ariel Halvey
drhalevy@gmail.com

Extended author information available on the last page of the article

(multiple choice) and an oral exam in order to graduate any residency program.

During the first years, the examinations were held at different surgical wards each time, based on the hosting surgery departments. The exam consisted of 3 committees, each of three members. The candidate was asked to perform one supervised physical examination during the assessment. The examination covered any surgical subspecialty, and the candidate had to master an in-depth knowledge of each surgical field in order to pass (i.e., vascular, plastic, pediatric, thoracic non-cardiac surgery, wound healing, burns, and urology). The cases presented to the candidate were random and based on the examiner's case log.

Since 1990, the examination has been limited to general surgery cases only, and subspecialities cases have been excluded. The examinations were, for the first time, held at a central location. Examiners were instructed to present relevant imaging, which had become an integral part of the exam. During this period, the exams lacked standardization as different candidates were asked about different cases, either the same or different examiners. Heterogeneity prevailed. The passing score was defined at 60%. At the end of the test, all examiners held a consultation, openly discussing all candidates. Discussions were held especially for candidates who did not receive passing scores by all stations members as well as candidates who did not successfully manage cases that were considered mandatory to pass, mainly trauma scenarios.

Since 2000, there has been an ongoing process of examinations' standardization which included the development of a questions bank — predefined scenarios with set questions. Some surgical fields have been abandoned and omitted (i.e., soft tissues sarcomas, malignant melanoma, endocrine surgery, peri-operative surgical management) in the process, mainly due to the lack of exposure of residents in some wards to these pathologies.

In 2018, a new exam concept was formulated by the authors of this study, aiming at improving standardization and extending the examined surgical fields to all major surgical domains. The main competencies to be assessed are clinical judgment and decision making, imaging interpretation, management of pre- and intra-operative surgical dilemmas, complications management, operative note writing, and to a lesser degree, theoretical knowledge. Questions bank has been discontinued, and new scenarios are written for each exam. The cases for the exam are written and standardized at a yearly meeting. Departments lacking all surgical domains were encouraged to send their residents to other surgical departments in order to improve exposure to specialized surgical fields.

All exams are held at a centralized location. Examiners are blinded to the hospital of the candidates. The exam currently consists of 12 stations of two examiners, each focusing on a different field:

- Written operative note: a single scenario is presented. The candidate is instructed to decide upon the procedure chosen and to write a full operative note, including perioperative care, and explain the rationale of the procedure he plans to perform. The note should include, in addition, consent process, with a summary of benefits and risks of the planned procedure.
- Breast surgery
- Endocrine surgery
- Metabolic surgery
- Trauma
- Acute care surgery
- Colorectal surgery
- Hepatobiliary and pancreatic surgery
- Abdominal wall surgery
- Surgical oncology (focusing on soft tissue sarcoma, malignant melanoma, cytoreductive surgery and surgical oncology and palliative surgery dilemmas)
- Esophago-gastric (foregut) and small bowel surgery
- Peri-operative care (focusing on physiology, pathophysiology, nutritional support, fluids and electrolytes imbalance, acid–base balance, coagulopathy in the surgical patient, respiratory and hemodynamic support)

Each scenario is constructed as a computerized slideshow presentation. The candidate is presented with a slide by slide as the scenario progresses. The examiners are presented with mandatory discussion points for each slide.

Each station is 25 min long and presents one pre-defined and well-constructed scenario to the candidate. The scenario presented in each station is based on actual patient cases of the relevant clinical field, modified to the examination's format: case history, laboratory data, imaging interpretation, pathological report interpretation, clinical and surgical dilemma, and complications management. All cases are written by experts in the relevant clinical field. Cases are sent pre-hand to the exam committee, where they are validated and standardized to match the exam's format. Turn-over time between stations is 5 min, and candidates get 15 min break during the exam. Candidates' cellphones are collected before exam commences, and are constantly monitored so they cannot discuss exam questions. The total duration of the exam is 320 min. Each candidate is presented the same scenario and asked the same set of questions.

Each station consists of two experts in the field (the perioperative care station consists of ICU specialists), who fill out an individual evaluation form, elaborating candidate's achievements, with a fail/pass/pass with distinction score. The examiners are instructed not to discuss their evaluation with each other to maximize objectivity.

As duplicate stations are set, depending of the number candidates, all examiners of the same field meet and discuss their scenario for the exam. This is to ensure standardization between stations. Overall, each candidate is evaluated by 24 examiners, minimizing bias in the process of final evaluation. Examiners who are familiar with a candidate are temporarily replaced. Each candidate is called and informed of his final score at the end of the exam. In case of failure, the candidate may appeal and receive a detailed report of the exam outcome.

Examiners undergo a preparation course aimed at teaching the process of trainee evaluation. Preparation courses are currently mandatory to all candidates before their first exam, and are conducted by one of the exam committee members. The course includes assessment tools, tools to maximize exam objectivity, interacting with different types of candidates, and simulated examination scenarios. These courses were in practice for many years before the institution of the current examination format, but were not mandatory. Novice examiners are preferably teamed with an experienced examiner.

Resources and Costs

None of the examiners or members of the examination committee are being reimbursed for the examination day or exam preparation. It is entirely voluntary. Examination housing costs and administrative support are covered by the Israeli Medical Association. A ratio of 1:3 candidates to examiners is required for exam execution, which means a large number of examiners, all practicing surgeons, are necessary during exam day. This impacts the volume of elective case-load in many Israeli surgical departments.

Methods

Between 05/2018 and 11/2020, two exams per year were conducted with the current exam concept.

Candidates were evaluated individually by each of the two examiners in each of the 11 stations and received a pass or fail grade. Operative notes were graded by two independent examiners. Examination's passing score was defined as 80% — i.e., 19/24 positive evaluations. Pass with distinction was defined as passing all stations and receiving at least 10 pass-with-distinction evaluations.

At the end of the examination, examiners and candidates were asked to complete a questionnaire regarding the examination's process and quality (scale 1–5, questions and mean results are shown in Tables 1 and 2). All questionnaires were blinded. After the last station, a text message containing a link to the online questionnaire was sent the examiners. Questionnaires were filled by hand by the candidates immediately after the last station and before candidate being informed of the examination's results.

Table 1 Individual questions scores

Topic	Pass rate (%)*	Pass with distinction rate (%)*	Fail rate (%)*
Acute care surgery	86.4 ± 4.2	18 ± 7.9	13.6 ± 4.2
Colorectal surgery	84 ± 8	15.7 ± 8.5	16 ± 8.1
Esophago-gastric surgery	84.6 ± 7.2	16.8 ± 4.4	15.4 ± 7.2
Metabolic surgery	83.8 ± 4.9	14.3 ± 4	14.6 ± 13.4
Peri-operative care	83.4 ± 7	17.8 ± 5	16.6 ± 7
Operative note	83.1 ± 9.3	15.3 ± 13.8	15 ± 9.1
Trauma	82.8 ± 3.3	15.6 ± 9.1	17.2 ± 3.4
Endocrine surgery	79.7 ± 9.9	23.3 ± 14.5	20.3 ± 10
Breast surgery	78.7 ± 9.3	25.3 ± 6.7	20.4 ± 10
Hepatobiliary surgery	77.8 ± 12.8	19 ± 6.4	22.2 ± 12.9
Abdominal wall surgery	71.2 ± 16.7	22.4 ± 10.1	28.8 ± 16.9
Surgical oncology	68.3 ± 13.3	18.7 ± 11	31.7 ± 13.3

*Mean

Results

A total of 142 residents attended six oral board examinations between 05/2018 and 11/2020. Mean pass rate was 76.6 ± 9.5% (range — 63.2–87%).

Questions with overall highest mean pass rates were acute care surgery (86.6 ± 4.8%), foregut surgery (84.6 ± 7.6%), and colorectal surgery (84 ± 8.1%). Questions with the highest rate of pass with distinction rate were breast surgery (23.3 ± 6.7%) and abdominal wall surgery (22.4 ± 10%). Questions with the highest mean fail rates were surgical oncology (31.7 ± 13.3%) and abdominal wall surgery (28.8 ± 16.9%; Table 1).

No case of COVID-19 exposure was recorded during the examination process and thereafter.

Candidates' Feedback

A total of 55 candidates responded to the questionnaires. Highest graded topics were as follows: did the examiners treat you in an appropriate manner (4.08 ± 1.17); examination's organization and logistics (3.94 ± 0.8) were the cases presented in a clear manner (3.92 ± 1.05; Table 2).

All candidates reported adequate length of time of the entire exam and of each station. One candidate recommended to decrease the total number of the cases (stations)

Table 2 Candidates' feedback

Topic	Mean score*
Did the examiners treat you in an appropriate manner	4.08 ± 1.05
Examination's logistics	3.94 ± 0.6
Were the cases presented in a clear manner	3.92 ± 1.17
Cases presented reflect the daily work of an attending surgeon	3.68 ± 1.31
Fairness of the examination	3.66 ± 1.34
Cases presented reflect topic you were exposed to as a resident	3.64 ± 0.87
Were you able to express your knowledge during the examination	3.35 ± 0.83

*Scale of 1–5: 1 — least appropriate, 5 — most appropriate

of the examination, as he felt the exam was too long. Half of the candidates reported adequate difficulty level of the examination, and the other half as too high. Due to anonymity of the questionnaires, the examination committee was unable to correlate candidates' final scores with their subjective assessment of the examination's difficulty level. Examination's organization and logistics during COVID-19 pandemic were scored as high as 3.93, and verbally commended. This was comparable to previous examination's scores.

Examiners' Feedback

A total of 162 examiners responded to the questionnaires. Highest graded topics were as follows: standardization of the exam (4.45 ± 0.63), fairness of the exam (4.44 ± 0.67), and whether the presented cases reflect the daily work of an attending surgeon (4.35 ± 0.87; Table 3).

Most (91.9%) of the examiners reported the duration of the examination as adequate, 71.2% of the examiners reported the difficulty level of the examination as adequate, and 22% reported it as too low. Six (3.6%) examiners evaluated the examination as too long, and reported difficulty in maintaining adequate level of concentration when repeatedly asking the same set of questions. Three examiners recommended changing the examination's format so that each station presents two cases instead of one or adding more scenarios.

Discussion

The primary goal of medical education is clinical performance; however, there is no consensus regarding the best evaluation method [1]. The problem is further complicated in the surgical field — should we test surgical knowledge alone or should surgical abilities be tested as well. If so, what is the best way to do it? No ideal exam model has been suggested in general surgery for trainee evaluation. Oral exams are challenging and bear limitations, but allow assessment of trainees' thinking process. Several models exist for oral trainee evaluation, with the Objective Structured Clinical Examination (OSCE) model [2, 3] being one of the first developed during early 1970s, aiming at better standardization. In the OSCE model, the exam is comprised of a rigid structure of questions with expected responses by which the candidate is evaluated. The model presented here slightly differs from the OSCE model [2], as each scenario is openly discussed with the candidate, allowing a better evaluation of the candidate's thinking process and decision making skills.

The format of the exam is constantly evaluated in a process based on exam's results and feedback analysis of the examiners and examines. The first exam performed in the current exam concept was more similar to the classic OSCE — each scenario was comprised of several predefined questions. Each question had "right" and "wrong" responses for which the candidate either received or lost a point. Even though the exam had highest pass rate (87%) compared to the following exams, many examiners commented that this format was inadequate, as it lacked the ability to evaluate

Table 3 Examiners' feedback

Topic	Mean score*
Inter-candidate standardization	4.45 ± 0.63
Fairness of the examination	4.44 ± 0.67
Cases presented reflect the daily work of an attending surgeon	4.35 ± 0.87
Suitability of the examination format for oral board certification	4.14 ± 0.98
The evaluation form accurately reflects candidates' achievements	4.04 ± 0.82

*Scale of 1–5: 1 – least appropriate, 5 – most appropriate

the candidates' thinking process and decision-making skills. As these qualities are considered paramount for the board-certified surgeon, the exam format was further revised. Each scenario is constructed as a computerized presentation. The candidate is presented with a slide by slide show as the scenario progresses. The examiners are presented with mandatory discussion points for each slide. These include patient workup strategies, pre- and intra-operative dilemmas, clinical and radiological data integration, and complication management.

Surgical residency in Israel follows a broad curriculum, which is reflected in the clinical fields covered by the oral board. Blueprinting for such an exam is a challenging process. The main competencies assessed in the exam are clinical and surgical decision making, imaging interpretation, problem solving, and treatment planning. Because of the complexity of clinical competence evaluation, different tests are used during training. The use of multiple examiners across different cases improves the inter-rater reliability of the exam [4], as the average judgment of 24 examiners, each assessing the candidate on one question, produces a reliable test. The assessment of the consistency of the candidate's performance between cases is more challenging, as it greatly influenced by the experiences encountered during training. The cases presented are based on real patients, but are modified and standardized — case history, relevant laboratory values and imaging, clinical dilemma, pre- and intra-operative decision making, and complication management. The examiners are provided with mandatory discussion points to maximize reliability and objectivity. In addition, all examiners of duplicate stations participate in a pre-exam calibration meeting.

Oral exam is an essential part of the certification process in other countries as well. In the UK, candidates undergo two examinations as a mandatory part of their certification [5] — section 1 is a multiple-choice exam designed to test the application of knowledge and clinical reasoning. Section 2 comprises the clinical component of the examination. It consists of a series of interviews on clinical topics — either scenario or patient-based. The competencies assessed are of knowledge, clinical interpretation, decision-making, clinical judgment, and professionalism. The candidate is independently marked by examiners working in pairs, but with reference to the standard agreed at a pre-exam calibration meeting. Similar to Israel, the cases presented include predefined discussion points for the examiners [6]. The American Board of Surgery (ABS) mandates a written multi-choice exam (qualifying exam) and an oral exam (certifying exam) [7]. The oral exam consists of 3 consecutive 30-min stations, each of two examiners. The examiners are an ABS director and an experienced ABS diplomate, from the regional medical community. Four cases are presented to the candidate during each 30-min station. Similar to Israel, the candidates are evaluated of 12 cases, but each case is shorter. The candidates are evaluated by 6 examiners. In Israel, the candidates are evaluated by 24 examiners on 25-min cases, which provide a broad evaluation platform.

The fields in which candidates scored highest, i.e., breast surgery, acute care surgery, and colorectal surgery, were the fields in which most candidates received extensive exposure during their training. Fields in which candidates scored poorly are addressed in the preparation course for the oral board exam. Negative feedback for the exams mainly focused on scenarios in areas in which candidates had little exposure — i.e., esophageal surgery, complex hepatobiliary surgery, and surgical oncology — as most candidates had little exposure to these pathologies, especially in smaller medical centers.

The current examination format is extremely challenging to maintain — the number of examiners needed is a 1:3 ratio to the number of candidates. This has a significant impact on the daily work of the general surgery departments in Israel, as it is forced to reduce elective case-load on the day of the exam.

COVID-19 pandemic presented significant challenges for holding large events. As oral board examinations are critical and mandatory part of residents' training and evaluation, special measures had to be taken to ensure both safe and effective examination, in order not to delay their certification process, which, in turn, would affect the number of senior general surgeons in Israel. Prodigious efforts were made to set-up the examination process as similar as possible to previous exams, before COVID-19 pandemic, as opposed to other models, as suggested by Lara et al. [8]. They reported their experience with teleconference OSCE ("TeleOSCE") for clerkship students at the completion of their pediatric rotation. Forty-nine students were tested over 3½ days, with all interactions done by video-conferencing (student-simulated patient-observer). They reported comparable exam quality and reliability with live-OSCE. A model of tele-examination was considered by the Israeli examination committee, but was abandoned. The committee felt face-to-face interview has significant added-value compared to video-conference facilitated exam.

The model presented herein is similar to that reported by Boursicot et al. [9]. The authors reported their experience in OSCE performance for medical school graduates during COVID-19 pandemic. Similar to our practice, briefings were video-conference facilitated, and social distancing was strictly kept throughout the examination process. Our examination slightly differs from classical OSCE, as each case is openly discussed with the candidate, allowing a better evaluation of the candidate's thinking and decision-making process.

Conclusions

Oral exams are challenging and bear limitations, but properly constructed exams allow good evaluation of the trainees' thinking process and decision-making skills, which are paramount to the board-certified general surgeon, without compromising exam's integrity and standardization. The importance of this report is the description of our novel model for the oral board exam.

Declarations

Competing Interest The authors declare no competing interests.

References

1. Colliver JA, Swartz MH (1997) Assessing clinical performance with standardized patients. JAMA 278(9):790–791
2. Sloan DA, Donnelly MB, Schwartz RW, Strodel WE (1995) The objective structured clinical examination. The new gold standard for evaluating postgraduate clinical performance. Ann Surg 222(6):735–42
3. Setyonugroho W, Kennedy KM, Kropmans TJ (2015) Reliability and validity of OSCE checklists used to assess the communication skills of undergraduate medical students: A systematic review. Patient Educ Couns S0738-3991(15):00277–3
4. Wass V, Van der Vleuten C, Shatzer J, Jones R (2001) Assessment of clinical competence. Lancet 357(9260):945–949
5. Program ISC. General Surgery Curriculum [Web site]. ISCP web site2021 [Available from: https://www.iscp.ac.uk/iscp/curriculum/general-surgery-curriculum/5-programme-of-assessment/
6. Examinations TJCOI. General Surgery Curriculum [Web page]. JCIE; 2015 [Available from: https://www.jcie.org.uk/content/content.aspx?ID=14
7. Surgery ABo. General Surgery Certifying Exam [Web page]. 2022 [Available from: https://www.absurgery.org/default.jsp?certcehome
8. Lara S, Foster CW, Hawks M, Montgomery M (2020) Remote assessment of clinical skills during COVID-19: a virtual, high-stakes, summative pediatric objective structured clinical examination. Acad Pediatr 20(6):760–1. https://doi.org/10.1016/j.acap.2020.05.029
9. Boursicot K, Kemp S, Ong TH, Wijaya L, Goh SH, Freeman K et al (2020) Conducting a high-stakes OSCE in a COVID-19 environment. MedEdPublish 9(54). https://doi.org/10.15694/mep.2020.000054.1

Authors and Affiliations

Guy Pines[1,2] · Ofir Ben-Ishay[3] · David Hochstein[3] · Guy Lahat[4] · Gil Ohana[5] · Gilbert Sebbag[6] · Yoram Kluger[3] · Ariel Halvey[7]

1 Faculty of Medicine, Hebrew University of Jerusalem, Jerusalem, Israel

2 Department of Thoracic Surgery, Kaplan Medical Center, Jerusalem, Israel

3 Rambam Medical Center and the Technion Medical School, Haifa, Israel

4 Tel Aviv Medical Center and Sakler Medical School, Tel Aviv, Israel

5 Barzilay Medical Center, Ashkelon and the Ben-Gurion University Medical School, Beer-Sheva, Israel

6 Soroka Medical Center and the Ben-Gurion University Medical School, Beer-Sheva, Israel

7 Emeritus Professor of Surgery, Sakler Medical School, Tel Aviv, Israel