The American Board of Obstetrics and Gynecology’s remote certifying examination: successes and challenges

Pooja Shivraj, MS, PhD; Rajat Chadha, PhD; Deana Dynis, EdD; Kirk Diepenbrock, MS; Barbara L. Hoffman, MD; Susan Ramin, MD; George D. Wendel, Jr. MD

In response to the COVID-19 pandemic, the American Board of Obstetrics and Gynecology canceled the 2020 in-person subspecialty certifying examinations and developed remote administration of 4 subspecialty certifying examinations in 2021 for both examiners and candidates. Because of the continued risks of the COVID-19 pandemic, the 2021 specialty certifying examinations and the 2022 subspecialty certifying examinations were also administered remotely for candidates. For these examinations, examiners participated remotely in 2021 and were at the American Board of Obstetrics and Gynecology testing center in 2022. Overall, the American Board of Obstetrics and Gynecology remote certifying examinations have been well-received by candidates and examiners according to posttest survey data. Candidate performance has been comparable to that observed in the previous in-person examinations. In this review, we describe our implementation, process modifications, successes, and challenges with remote testing. During this process, the American Board of Medical Specialties approval was required, and the Standards for Educational and Psychological Testing served as our testing-industry guideline to ensure valid interpretation of scores and fairness to candidates.

Keywords: American Board of Obstetrics and Gynecology, assessment, certification, certifying examination, challenges, COVID-19, fairness, obstetrics and gynecology, remote, technology, testing, virtual

Introduction

The nation witnessed a change in many organizational processes in 2020 as a result of the SARS-COV-2 (COVID-19) pandemic. Specific to the world of testing and educational measurement, examination administrations were either delayed or canceled, item writing workshops were held remotely, and testing and certification organizations sought to develop fair, safe, and long-term solutions to meet the needs of test-takers, academic volunteers, and other relevant stakeholders to keep them safe and engaged. In response to the pandemic, the American Board of Obstetrics and Gynecology (ABOG) canceled its face-to-face certifying (oral) examinations (CEs) for approximately 250 candidates in April 2020 across the 4 subspecialties that include: female pelvic medicine and reconstructive surgery, gynecologic oncology, maternal—fetal medicine, and reproductive endocrinology and infertility. By that time, only a handful of the 14 member boards of the American Board of Medical Specialties (ABMS) that offer oral examinations had initiated a plan to move the examinations to a virtual format, and ABOG was not prepared to administer the examination virtually.

To ensure a continued pathway to certification for the 2020 and 2021 candidates, ABOG rapidly developed a remote format for all subspecialty CEs for April 2021. During the year delay, several ABMS member boards had administered remote oral examinations and shared successes and failures to bolster a collaborative approach to help develop the ABOG remote CE. None of these ABMS boards had oral examination formats that included assessments based on personal practice similar to the ABOG CEs, and no board includes a thesis defense during their subspecialty CE administrations. This expert review shares our process of administering the specialty (general obstetrics and gynecology [OB/GYN]) and subspecialty CEs. We highlight the necessity for fairness and validity in testing, which is outlined in Standards for Educational and Psychological Testing, and discuss implications for administering CEs remotely in the future.

Process modification for successful administration

Since the beginning of the pandemic, ABOG has successfully administered the four 2021 subspecialty CEs, the (4-week) 2021 OB/GYN specialty CEs, and the four 2022 subspecialty CEs remotely. Specifically, 1491 specialty candidates in 2021 and 830 subspecialty candidates from 2021 and 2022 combined took their CEs. For this, process modifications were required to ensure the fidelity and security of the CE format and content.

Traditionally, ABOG’s specialty and subspecialty CEs are 3-hour tests composed of three 1-hour sessions that assess different elements of OB/GYN practice.
Questions regarding simulated patient scenarios (structured cases) and management of the candidate's personal practice case list are answered by the candidate.\textsuperscript{5,6} Each of the 3 sessions is administered by 2 unique examiners, and this structure provides assessment by 6 examiners per candidate. When administered in-person, the CEs are held in a standardized setting at the ABOG testing center in Dallas. During an examination, the candidate remains in the same room for the 3-hour test and is joined by the respective paired examiners for each 1-hour session. For many candidates, this requires air travel, hotel expenses, and time away from productive clinical practice and home.

For the ABOG remote test administration, ABOG used Zoom—a collaborative, cloud-based video conferencing software.\textsuperscript{7} Each candidate was at a nonstandardized home or work site, and they were assigned to a single 3-hour Zoom appointment. For examination security, a trained virtual proctor, hired from an external procuring company and allowed to monitor candidates from any location of choice, stayed in the candidate's Zoom room to monitor activity. Training for proctors included how to register a candidate, and documentation of minor/major technical and security concerns, as defined by ABOG. For example, loss of connectivity vs a drop-off from the call might constitute the difference between minor and major technical issues, whereas finding an app open during registration vs finding a recording device during the exam might constitute the difference between a minor vs major security issue. In contrast, examiner pairs received several Zoom appointments, which moved them from one Zoom room to another. To coordinate this, links were sent to examiners and candidates well in advance, and an external vendor was used for this process.

As another modification, all cases were embedded within Portable Document Format (or PDF) files and shared with the candidate visually in real time through the Zoom links. Computer displays and bandwidth variation among candidates could not be controlled. Thus, any forms of clinical videos were excluded from the structured cases to ensure that examination fidelity was not diminished. However, candidates were required to have an adequate internet bandwidth of 50 Mbps after their technical checks and by testing their internet speeds on fast.com.

Of resource allocations, in-person CEs require significant physical space at the ABOG testing center. In contrast, remote CEs required a substantial number of proctors, information technology (IT) resources and IT experts, and robust ABOG staff coordination and communication. Proctors were recruited and underwent standardized training to record technical- or security-related incidents. Before the start of an examination, proctors conducted a 360-degree room scan of the candidates' surroundings to ensure that notes or electronic devices were absent, and they also monitored the examination in real time for security-related incidents.

For CE administration, both candidates and examiners required training and practice regarding examination-day expectations and processes. Webinars and manuals were created, and both candidates and examiners prepared before their CE date to familiarize themselves with Zoom links and screen sharing. Attestations on completion of training were required from candidates and examiners to ensure that CE delivery would be more standardized and seamless. One-on-one virtual technical checks were conducted by ABOG on every machine (candidates and examiners) to confirm that every participant’s equipment met the minimum technical requirements.

Remote testing successes
ABOG successfully administered 100% of scheduled CEs on their assigned day for all 2021 subspecialty CEs. Despite connection difficulties, we successfully administered 100% of scheduled examinations for the 2021 specialty and 2022 subspecialty CEs within the assigned week. Only 4 candidates took their test on a different day/time than originally scheduled, but still within their originally assigned week, because of significant disruption issues in connectivity.

Two-proportion z-tests showed no significant differences in pass rates between candidates taking each of the remote CEs and candidates taking the respective previous in-person 2019 subspecialty CEs. This was true for first-time takers and for all test takers (first and repeat takers). Pass rates can be found on the ABOG website.\textsuperscript{8}

Proctors reported no major security concerns within their sessions. This was also confirmed by examiners during their grading.

Posttest surveys were sent after every administration to both candidates and examiners. Across the 4 subspecialty examinations, 86.7% of candidates regarded their ability to communicate their knowledge, skills, and abilities to be similar to when they took their previous in-person specialty CEs. A total of 95.1% of candidates were satisfied with the convenience of remote examination, and 95% of examiners across all CEs considered that candidates’ outcomes would have likely been the same regardless of in-person or remote administration. Consistent survey results were observed across all 4 subspecialties.

Remote testing challenges
Most examination sessions had no connectivity or audio/video disruptions. However, up to 1.1% of either candidates or examiners experienced ≥1 of these technical issues, which disrupted testing to some degree.

ABOG staff faced logistical challenges in scheduling candidates and examiners to accommodate for all US time zones. All departments at ABOG worked additional 4 to 6 hours daily during CE weeks to ensure examination fidelity. Similarly, candidates and examiners in certain time zones also experienced excessive scheduling burden. This was especially true for those in Hawaii Standard Time.

In posttest surveys, many candidates and examiners reported some difficulty with the remote administration: 41.7% of candidates reported anxiety in taking the CEs remotely compared with in-person, whereas 40.4% of candidates reported their anxiety as about the same, and 27.7% perceived communicating with
examiners to be more difficult remotely. Of the examiners, 53.6% also reported anxiety when administering CEs using the Zoom platform compared with when they administered previous in-person examinations, whereas 42.1% reported their anxiety to be about the same, and 41.4% described difficulty in reading candidates’ facial expressions and body language remotely.

Finally, with the pushback against remote proctoring, petition campaigns, and lawsuits about violation of privacy, balancing the protection of candidates’ privacy and the efforts to maintain integrity of the examination and deter cheating will likely remain a challenge in the near future of remote testing.

Testing standards

The Standards for Educational and Psychological Testing provide criteria for the development, administration, and scoring of tests, and guidelines for the interpretation and use of the obtained scores. To maximize fairness, the Fairness in Testing standards delineate different safeguards that developers can use to protect test-takers during test development and administration. One is the elimination of sources of bias that are not relevant to measuring the knowledge, skills, and judgment needed to be a competent OB/GYN specialist. These sources of bias are known as "construct-irrelevant variance" and may affect examination fairness and threaten the validity of pass–fail decisions. Construct irrelevance is the extent to which scores are affected by processes external to the test’s intended purpose and that are not part of the test’s construct, and it may take on many forms in the context of this study.

For an ABOG remote CE, these variances may include anxiety with technology (on the candidate or examiner end) or technical disruptions, poor image quality, uneven examination duration, or differing hardware or bandwidth among candidates. Standard 3.4 states that “[t]he environment should furnish reasonable comfort with minimal distractions to avoid construct-irrelevant variance." Examples include, but are not limited to, background noise, any disruptions in the testing area, poor lighting, and equipment failure during testing. A standardized testing environment is nearly impossible with remote testing, and this burden falls to the candidate. Pets at home, external construction, and announcements over hospital public address systems are common, unforeseen disruptions experienced during the ABOG remote CEs. Moreover, bias may be introduced by the environment. A candidate testing in a university office may be viewed more favorably than one testing in their home office or den. These sources of construct-irrelevant variance can be largely mitigated by a centralized testing center.

Last, Standard 6.6 discusses eliminating opportunities for cheating or attaining scores by fraud. Minimizing the possibility of cheating and detecting security breaches requires extra effort and vigilance during any remote CE. This necessitates additional visual room sweeps, proctoring, candidate pocket checks, and desktop scans, among others, using proctors. In addition to using proctors during the examination, an outside vendor assisted ABOG before, during, and after the week of the CEs by providing web patrolling services to search for leaked or harvested questions. The vendor did not find any of the test content related to the examinations discussed herein during its web patrolling.

Discussion

Determining the future of ABOG’s examination process will weigh successes and disadvantages. Of the advantages, candidates reported that remote administration was more convenient, and many appreciated not having to travel. Second, overall pass rates did not differ from those of previous years. This suggests that CE fidelity was preserved despite remote administration.

However, the challenges of remote administration to fairness and examination security are substantial. First, variability in technical equipment, testing environment, and even communication quality could lead to systematic biases in rating of candidates by examiners. Second, technical disruptions in small, once-a-year examinations are insufficient to warrant rescheduling, which can delay certification. Because many hospitals require board certification, rescheduling could pose unintended consequences for employment. Third, anxiety reported by both candidates and examiners regarding virtual examinations cannot be understated. Candidates’ anxiety seemed similar to that experienced with in-person examinations, and thus may stem from anxiety related to examinations in general. However, examiners reported differences in anxiety in remote vs in-person examinations, which could be related to administering the examination through technology. Finally, although overall pass rates on the remote CE may not be different from those of in-person examinations from previous years, ABOG witnessed multiple candidates who had technical failures and/or disruptions during the examination, many of which could have had a (negative or
positive) impact on their final outcomes.

During the peak of the pandemic, many organizations were successful in administering remote oral examinations, and some are considering continuing with this model.\textsuperscript{10−12} However, it is important to note that most of ABMS boards’ CEs are not as complex as the ABOG assessments that include examinations of personal practice cases and subspecialty theses. As with any remote examination, the major limitations of the ABOG administration were that security breaches could not always be detected, and construct-irrelevant sources may have played a role in candidates’ performance and outcome. With contemporary sophisticated recording devices, theft of copyrighted examination content is a continual threat. These environmental biases and potential security hazards are largely mitigated in a testing center.

The 2022 ABOG specialty CEs are being administered to candidates remotely, and the examiners will be at the ABOG testing center. As the pandemic comes to an end, ABOG will rely on the Standards to support our certification decisions. The testing challenges of construct-irrelevant variance, technological variances and constraints across candidates, potential security breaches, environmental biases, and administrative burden outweigh the benefits of candidate convenience of taking examinations from home, especially for candidates who face these challenges. To ensure testing fairness, validity, and security, there is a strong case to be made for future CEs to be administered in person. Our intent is to return to examination administration at our testing center in 2023.

REFERENCES
1. Radin J, Korba C. COVID-19 as catalyst: the future of work and the workplace in health care. 2020. Available at: https://www2.deloitte.com/us/en/insights/industry/health-care/health-care-workforce-trends.html. Accessed September 3, 2022.
2. Bartley GB. COVID-19 and the American Board of Ophthalmology: when the best-laid plans go awry. Ophthalmology 2020;127:e55–4.
3. American Board of Medical Specialties (ABMS). ABMS member boards oral exams go virtual. 2022. Available at: https://www.abms.org/news-events/abms-member-boards-oral-exams-go-virtual/. Accessed September 3, 2022.
4. American Educational Research Association. American Psychological Association. National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: American Educational Research Association; 2014.
5. American Board of Obstetrics and Gynecology (ABOG). Subspecialty certification bulletins. 2022. Available at: https://www.abog.org/subspecialty-certification/bulletins. Accessed September 3, 2022.
6. American Board of Obstetrics and Gynecology (ABOG). Specialty certification bulletins. 2022. Available at: https://www.abog.org/specialty-certification/bulletins. Accessed September 3, 2022.
7. Zoom video conferencing software. 2022. Available at: https://www.zoom.us/. Accessed September 26, 2022.
8. American Board of Obstetrics and Gynecology (ABOG) Website. 2022. Available at: https://www.abog.org/. Accessed September 4, 2022.
9. Young JR. College’s use of exam proctoring software to ‘scan’ rooms violated privacy rights, judge finds. 2022. Available at: https://www.edsurge.com/news/2022-08-25-college-s-use-of-exam-proctoring-software-to-scan-rooms-violated-privacy-rights-judge-finds. Accessed November 12, 2022.
10. Bartley GB, Schnabel SD, Comber BA, Gedde SJ. The American Board of Ophthalmology virtual oral examination: crisis catalyzing innovation. Ophthalmology 2021;128:1669–71.
11. Kossoff L. Will virtual oral board exams remain an option after the COVID-19 pandemic? 2021. Available at: https://www.entropy.org/article/will-virtual-oral-board-exams-remain-an-option-after-the-covid-19-pandemic2/?singlepage=1. Accessed November 4, 2021.
12. Chen H, Tseng JF, Chaer R, et al. Outcomes of the first virtual general surgery certifying exam of the American Board of Surgery. Ann Surg 2021;274:467–72.