SCIENTIFIC CULTURE

www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 2.1, (2026), pp. 4225-4233
Open Access. Online & Print

DOI: 10.5281/zenodo. 18936184

REMOTE CLINICAL SKILLS ASSESSMENT FOR MEDICAL PROFESSIONAL EXAMINATIONS: DURING AND BEYOND THE PANDEMIC

Avgis Hadjipapas[1*], Valentina Ionova[1], Amer Salim[2], Dionysis Vaidakis[1], Kevork Hopayian[1] and Ourania Kolokotroni[1]

[1]University of Nicosia Medical School, Nicosia, Cyprus
[2]Falcon Road Medical Centre, London, UK
[1]Current address: Department of Nursing, School of Health Sciences, Cyprus University of Technology, Limassol, Cyprus

Received: 01/01/2026
Accepted: 02/03/2026

Corresponding Author: Avgis Hadjipapas (hadjipapas.a@unic.ac.cy)

ABSTRACT

To overcome the challenges of the COVID19 pandemic, we developed a remotely-delivered, online clinical skills examination using purpose-built software. Here we outline and evaluate the reliability and validity of the exam and consider its application beyond the pandemic. Building on an existing digital scoring and virtual assessment software, Qpercom developed a new tool with input and repeated pilot testing by our academic and administrative examinations team. This tool also had additional features allowing for reliable (a) video - conferencing virtual examination rooms, (b) automatic control of the rotation of candidates across virtual rooms (c) a 'control room' for exam administration and (d) real time display of examination cards with physical examination and investigation findings. The new tool was used over three examination diets and the validity and reliability of the resultant examinations were assessed. The validity and reliability of the remote exams were not significantly different to those of face-to-face assessments. Content validity, was very similar with the exception of the inability to test physical examination. Regarding reliability, low Cronbach's alpha values were observed for some remote examinations when variability between student marks was also low explaining these alpha values. However, Standard error of Measurement (SEM) estimates, which are more reliable in the setting of small numbers examinations such as the ones described here, compared favourably to previous face to face examinations. Thus, remote, online delivery of clinical skills examinations using purpose-built software is feasible, valid and reliable and may represent a favourable option beyond the pandemic.

KEYWORDS: Remote Assessment, Pandemic, Clinical Skills Examination, Medical Professional Examinations, Healthcare Practitioners, OSCE Software.

1. INTRODUCTION

1.1. Objective

The aim of this paper is to outline the development of a remotely-delivered, online clinical skills examination using purpose-built software and provide its reliability and validity for its wider use in clinical skills assessment.

This type of examination was developed in response to the unprecedented challenges to health care education during the COVID-19 pandemic. Educators involved in the delivery of clinical courses like Medicine and Nursing had to rapidly develop alternative teaching solutions adapted to the new environment, since graduates of these programs were urgently needed to increase the healthcare workforce on the ground(Almarzooq et al., 2020; Hilburg et al., 2020). This led to a large variability in the approaches adopted to teaching clinical and communications skills, which had traditionally taken place face to face in teaching labs(Servin-Rojas et al., 2022; Wallace et al., 2021). Concerns were raised in regards to the impact of the new learning methods on the competency level of the future healthcare workforce, especially healthcare students in the final year of study and medical residents(Hamamoto Filho et al., 2022). At the same time, continuing certification of students and trainees necessitated virtual alternatives to face to face assessment of clinical skills and several approaches were developed (Chan et al., 2023; Giri & Stewart, 2023).

The development of tools for the reliable assessment of clinical competency and skills for undergraduate students and graduate fellows was of immense importance. Despite the fact that accrediting bodies were providing flexibility to clinical programme directors as to how to assess clinical skills(Patrício et al., 2013), the use of new technological methods such as zoom, skype and Microsoft teams, not previously tested in these settings, was unprecedented but also challenging.

In Cyprus, when the pandemic started in 2020, strict lockdown measures and travel restrictions were imposed by the Cyprus government (Quattrocchi et al., 2020). Thus, there was a need to adapt the delivery of the assessment methods of the Cyprus MRCGP [INT] programme (International Membership of the Royal College of General Practitioners, UK) which is delivered by the University of Nicosia Medical School in Cyprus and is supported by an online Masters programme in Family Medicine (Hopayian et al., 2019). The Cyprus MRCGP [INT] assessment programme consists of a written examination (the applied knowledge test-AKT), a portfolio of workplace-based assessments (WPBA) and a clinical skills examination: Simulated Surgery (SS).

Whilst the AKT was delivered online with remote proctoring software and WPBA is assessed by video recording of skills and consultations, the most challenging component to deliver in the pandemic, was the SS examination. The SS examination assesses clinical skills and competencies as per the objective structured clinical examinations (OSCEs) guiding principles (Patrício et al., 2013). The reliability and validity of OSCEs is dependent on a number of factors including the number of examination stations, the number of assessors and simulated patients and many more (Trejo-Mejía et al., 2016) The impact of introducing another variable such as the delivery of the exam with any online tool on examination reliability (but also partly validity) was unknown.

In order to address this need, we explored the option of utilizing and contributing to the development of a purpose-built software for remote clinical exams delivery, which allows for real time video assessments, purpose-built for OSCEs.

In the following sections of this paper, we describe the methodology followed for the implementation of the new tool in an online SS examination and we provide the preliminary results of the reliability and validity of its use in a series of remote, online exams.

2. METHODS

2.1. The Tool

The new tool developed by the Qpercom company was Qpercom Observe. While the actual software was developed by Qpercom, our assessment team provided substantial input and repeatedly tested the software while it was under development providing feedback to the company. The software had additional features required for delivery of online exams which include: (a) a reliable video -conferencing virtual room for student, simulated patient, examiner (and external observers if applicable) to meet, (b) automatic control of the rotation across 'stations' (virtual rooms) such that accurate and timely entry to the 'stations' of all participants is ensured 3) a 'control room' for examination administration staff, which allows real time monitoring and overview of the exam process , which in turn allows for real time trouble-shooting and problem solving 4) real time display to the student of examination cards showing physical examination and investigation findings.

Below we describe the setup of our examination process for the Cyprus MRCGP [INT] Simulated Surgery exam using the remote version of the

Qpercom Observe tool over 3 examination diets.

2.2. Setup

The setup is shown schematically in figure 1. The simulated patient and examiner are located in separate rooms on the premises of the university, while the student is located in their own home or practice. Administrative and technical support staff are also located on university premises, where they can intervene in case there are problems with the audio/visual or exam synchronization issues. External examiners or other observers can join the exam either remotely or on university premises. All participants connect to a remote, virtual room using Qpercom Observe software, where the examination takes place.

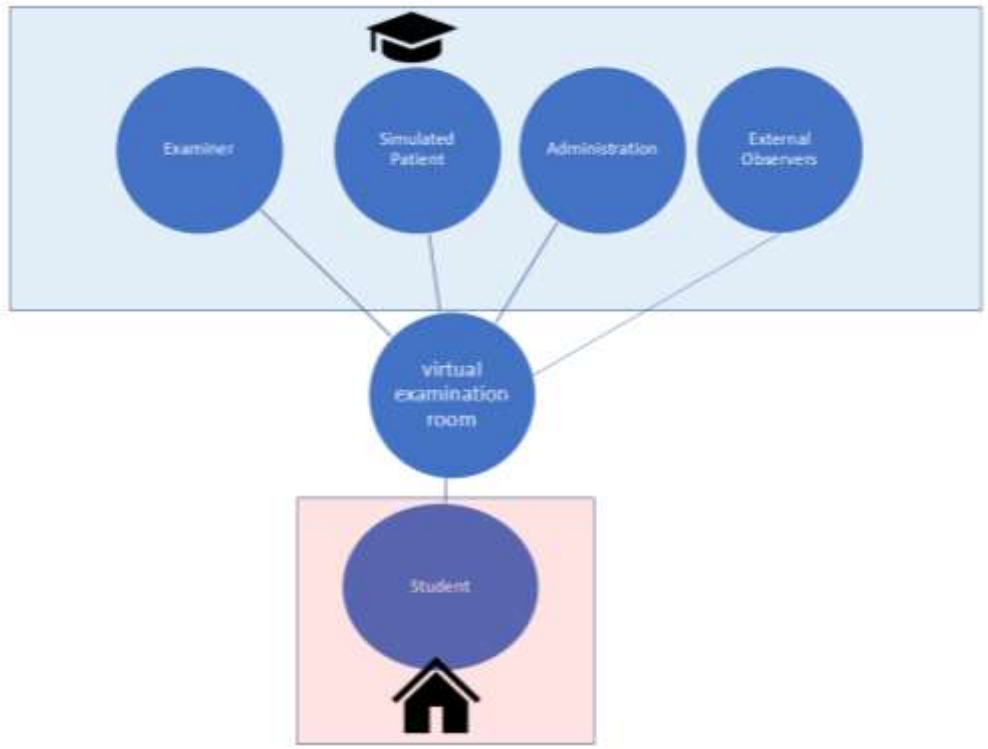

Figure 1: Setup of the Remote SS Examination
For an Explanation, See Main Text.

2.3. Examination Process and Administrative Control of the Remote Examination

First, examiners and simulated patients are briefed and their connectivity and audio-visual technical parameters are checked to ensure a proper connection between both parties. Then individual student briefings follow. The student briefing includes audio and visual checks and checking the internet browser used. At this stage, examination regulations are provided. Further notifications, such as stage of examination, are given throughout the session using features built-in in Qpercom software.

The examination process is made up of three stages (see figure 2): Reading Time, Assessment, Check Point/End of Station. These stages are repeated for each rotation/station examined. Reading Time: students read instructions prior to joining the examiner and actor in the virtual room. Assessment Stage: students enter the virtual room and the actual examination takes place. Check Point Stage: Once the station time has elapsed, students move from the virtual room and are placed in a virtual waiting room (the 'Check Point') ahead of the next station. The check point stage acts as a safety net because it allows for different station ending times within a circuit, which may result from any possible technical problems at any point in the circuit. At the end of the station each candidate is moved into a virtual waiting area until all candidates within the circuit have completed their station for the rotation. Next, the stations in the circuit are resynced, allowing the circuit administrator to initiate the process for the next rotation. Examiners and simulated patients remain in the same virtual room, while students are rotated between these rooms. End of Station: Candidates are informed of the end of the examination, asked to log out, and then released from the virtual examination room. Examiners are then

directed to select the next group of candidates to be assessed with the preceding steps/ stages repeated.

Thus, at all stages of the examination process, the purpose-built software provides functionalities, which enable full administrative control. This in turn, allows for effective management of any potential technical issues on a station or circuit basis. The ability for real-time notification and communication of exam administration and candidates, provides the means for an immediate response when troubleshooting without impacting the flow or order of stations within or across circuits.

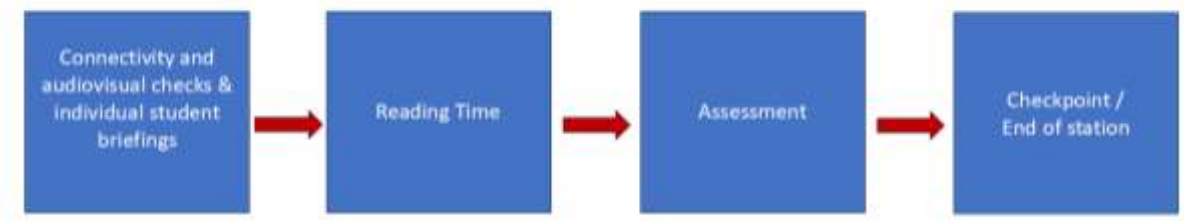

Figure 2: The Examination Process in the Virtual Room.
See Main Text for Explanation.

2.4. Sharing of Examination Findings and Tests between Examiner and Student in Real Time

An additional feature was required for the SS examination. This was the prompt sharing of content (examination and/or investigation findings) to the student at arbitrary time points in the consultation.

While in remote examinations one cannot test students' physical examination skills, diagnostic reasoning and result interpretation skills pertaining to the physical examination can still be tested. The student is presented with the examination findings in the form of a card or a picture only if the student has asked to perform the relevant physical examination. The student interprets the results in the overall context of the case and uses this information to arrive at a suspected diagnosis and formulate their management plan.

The examiner's ability to release examination cards permits evolution of the case giving the student the opportunity to demonstrate competency in results interpretation and formulation of a management plan. Implementing this feature initially presented a technical challenge in the software, with a fully-automated solution because the release of the cards is conditional upon the student's performance and cannot be pre-programmed.

Qpercom's solution was to have the cards as linked files but initially not visible ('inactive') on screen. The examiner can 'activate' these links to display the files immediately. Similarly, the examiner can withdraw the cards. This feature resolves the issues arising from the conditional nature of the cards and their variable timing.

2.5. Requirements for Administrative Support

Exam delivery in our case required a team of a minimum of five administrative officers per circuit implemented; three officers supported a participant group each (students, examiners, simulated patients). The fourth officer managed centrally the electronic examination delivery via the software. The above setup proved to be effective and efficient in handling technical and unforeseen circumstances. Coordination within the team is vital, and this requires a fifth officer who acts as a coordinator. The coordinator integrates information, decides on appropriate action (such as rescheduling students with connection problems if necessary) in a timely fashion. They also maintain a log of events and ensure that all students are examined fairly. We note that administrative support also depends on the number of students examined and examination circuits.

2.6. Technical Support

Having simulated patients, examiners, and administrators on one site facilitated the setup of hardware and ensured functioning across all workstations. It was important to test peripherals (cameras, microphones, keyboards etc) both in advance as well as during the assessment preparation stage. A guide to the software was provided to students, in order to set up their individual workstation setup according to the technical requirements. Students were also provided with a practice test of two SS stations to familiarize themselves with the process. On the days of the actual examination, two technical support officers

were on standby. Minor IT issues were managed within the Check Point stage of the examination. However more demanding issues, such as students' persistent connectivity issues were further investigated by the IT support team. In such cases a student's examination slot was rescheduled for a later time within the same examination day, in a so-called emergency exam session. This emergency session thus needs to be factored in exam planning, when developing the examination timetable.

3. ANALYSIS OF VALIDITY AND RELIABILITY

3.1. Validity

This refers to the extent to which the examination examines what it is meant to examine. In this paper, we have focussed on content validity, which essentially entails a comparison of blueprints of the SS examination for the remote versus the face-to-face versions of the exams. We also utilized the input of our RCGP International Development Advisor (IDA) at the time (Dr Amer Salim, author A.S.), who provided an objective, expert evaluation of the examination.

3.2. Reliability

We quantified the internal-consistency reliability facet using measures from classical test theory, namely Cronbach's alpha (*alpha*) and Standard Error of Measurement (*SEM*). We calculated *SEM* based on the equation below:

$$\boldsymbol{SEM = S_x\sqrt{1 - alpha}} \quad \textbf{(1)}$$

Where S_x is the standard deviation of the total SS % score across the cohort of students and ***alpha*** is Cronbach's alpha.

4. RESULTS

4.1. Rate of Successful Online Examination Attempts

We have run three SS examinations in partly-remote mode. From a total of 37 students, 12 had to be rescheduled for a different time of the day in the emergency sessions but all of these were finally successfully examined. Only two students could not be examined within the same examination diet. These students had severe problems with the technical requirements of their setup and/or insufficiently stable internet connections. These two students could not be examined in the same diet and had to be rescheduled for the next exam diet. In total, we have successfully examined 35/37 (95%) students in the same exam diet.

4.2. Exam Validity and Reliability

In terms of content validity, our coverage of key content and skills representative of the curriculum (the examination blueprint) was very similar to those prior to the pandemic with the key difference being the inability to test physical examination. In addition, our RCGP International Development Advisor (IDA), served as an external observer to the three remote examination diets and was satisfied that our remote SS version was not substantially different to the previous face to face version. Here we focussed on content validity, one of the important, classical types of validity. However, we must mention that newer theoretical frameworks challenge the classical view of validity. Messick (Messick, 1989) redefined validity as a unified concept and not separate types of validity (e.g. content, criterion, construct validity). This is an argument-based framework supported by multiple sources of evidence. In this framework, validity concerns the meaning and use of test scores. As an extension of this, Kane (M. Kane, 2009; M. T. Kane, 1992), redefined validation as an argument-based multi-stage approach, in which each inference must be logically and empirically maintained. Thus, while Messick emphasized types of evidence, Kane emphasized how evidence supports each step in the reasoning process. In our setting, of a very limited sample of observations (N of examinees), statistical evidence used in these frameworks would not be reliable enough to gauge validity/validation. In addition, we thought that the classical concept of content validity is very widely known among Medical Educators, hence we focussed on that.

With respect to exam reliability, we used Classical Test Theory (CCT) to estimate Cronbach's alpha coefficient (alpha) and Standard Error of Measurement (SEM). CTT assumes one source of error, which may not apply to an examination of clinical skills and Generalizability~~ation~~ Theory (GT), which partitions the error variance to~~in~~ different sources could be considered. We should note that our MRCGP [INT] exams are small number assessments (see N of students examined in table 1). GT, requires a higher N than CTT to partition the error variance into different sources, thus CTT may be the better choice in this setting. An additional reason, for using CTT is that it is very widely used and there is an abundance of software that can compute its outputs (alpha and SEM). GT usually is applied to complex designs; here we had a simple setting (a single circuit, single examiner and SP per station). Even in CTT, estimation of parametric statistics is employed. Because both alpha and SEM depend on parameter estimation, in a small N setting also alpha and SEM are not robust and should always be interpreted with caution. The estimates are shown in table 1. The

estimates for alpha were low for the July and September 2021 remote exams but greatly improved by 2022. However, when examining the table, it is striking that the low alpha values were observed when the variability between student marks (the standard deviation, SD) was also low and vice versa. Indeed, a least-squares quadratic fit (nonlinear correlation) confirmed that 84% of the variance in alpha values can be explained by SD variation. This, in turn, is due to estimation errors because of the small N in our assessment (see(Harvill, 1991)). However, from table 1 we note that the SEM is not affected to the same extent by the variability in student scores. This is likely due to the normalization intrinsic to its calculation (see equation 1 and(Harvill, 1991)). In fact, the SEM estimates for all the remote examinations do not deviate substantially (The Wilcoxon rank sum= 6, p= 0.0571) from those obtained in previous face to face examinations. In fact, if anything, a trend for a smaller SEM exists in favour of the remote examinations. Thus, in conclusion, we have not observed an inflation of examination-irrelevant error variance according to the behaviour of SEM, a measure which is more robust than alpha in the case of small numbers assessments.

Table 1: Cronbach's Alpha (alpha), Standard Deviation of Scores in % (SD) and Standard Error of Measurement (SEM, %) for Each Year the Exam Was Run. Note That Exams With High SD (Variability) Also Have High alpha and Vice Versa.

Year of SS exam	Mode of delivery	N of students completed exam	alpha	SD (%)	SEM (%)
2022	Remote	16	0.84	9.69	3.87
2021 September	Remote	9	0.54	6.61	4.48
2021 July (delayed for 2020)	Remote	13	0.69	6.90	3.84
2019	Face to face	8	0.82	12.54	5.36
2018	Face to face	6	0.91	15.28	4.71
2017	Face to face	10	0.74	9.82	4.97
2015	Face to face	9	0.94	22.18	5.59

5. DISCUSSION

We have presented the development of an exam delivery using purpose-built software for the partly-remote delivery of our MRCGP [INT] clinical skills assessment during the pandemic. We have described our experience of using this mode of delivery and shown that the examination is feasible, valid, reliable, and manageable both administratively and technically.

There have been studies of other forms of online OSCEs (Ryan et al., 2020; Saad et al., 2022; Shaban et al., 2021; Shehata et al., 2021) delivered by generic video conferencing software. An early paper established the feasibility of an online OSCE in a large cohort ((Ryan et al., 2020)). Ryan and colleagues used Zoom and MS teams video conferencing software to deliver the OSCE; however, they only delivered one station per day; thus, rotation complexities and synchronization was not addressed in that setup. Shehata and colleagues ((Shehata et al., 2021)) provided a step-by-step guide on how to run an OSCE using Zoom video conferencing and virtual Zoom rooms. However, they found that extensive administrative and technical support was required. Shaban and colleagues ((Shaban et al., 2021)) conducted online OSCEs using MS teams video conferencing software to support the video consultation in conjunction with locally-developed synchronization time-management software to support the rotation between stations. They found that the examination is feasible, acceptable to stakeholders, valid and reliable. The main difference to our approach is the semi-manual synchronization. Although the times at which rotations should occur is automated via a website, the transition process from station to station is still manual. This leaves some possibility for error when compared to our approach with purpose-built system integrating videoconferencing and transitions within the rotations. Saad and colleagues ((Saad et al., 2022)) analysed and summarized the experiences of three medical schools in Australia using Zoom; some of the

exams were fully online but there were also several hybrid versions. The study used qualitative methodology to investigate stakeholders' perceptions and also quality assurance. They concluded that the methods were fit for assessment in the setting of the pandemic, value for money and relatively fault-free. However, threats to validity (when compared to face-to-face OSCEs) were also identified, which according to the authors limited the scope of remote assessment. Nonetheless, the authors concluded that "Whilst virtual clinical assessments are limited in their scope of assessing clinical competency when compared with the traditional OSCE, their integration into programmes of assessment does, in fact, have significant potential". A recent review of the field summarized 20 remote OSCEs ((Kunutsor et al., 2022)). They concluded that the delivery of remote clinical assessments is possible, nonetheless several limitations and challenges are identified. The review also concludes that assessments are acceptable to both students and examiners, and that there is moderate agreement with face-to-face clinical assessments.

In the postgraduate setting, both the Royal College of General Practitioners, UK and the Royal Australian College of General Practitioners have adopted forms of remote assessment. The Royal College of General Practitioners, UK created the Recorded Consultation Assessment (RCA). A survey of RCA examiners found that 60% believed that the recorded consultations had enough information to make a judgement but 38% felt less confident about judgements and standard setting than in the face-to-face assessments. Case selection by candidates and lack of complexity were reported as problems ((Botan et al., 2023)). The Royal Australian College of General Practitioners' Remote Clinical Examination uses Zoom to connect students with simulated patients. The stations are pre-programmed break out rooms and examiners watch the consultation in real time.

At the time of writing this manuscript, we could not identify any evaluation of postgraduate membership exams which were conducted with purpose-built software. We believe that our evaluation will be useful, as now more Universities and Colleges may be moving in this direction. There are certain advantages of using purpose-built OSCE software. The duration of stations, circuits and sessions can be controlled independently, which allows for the duration of these to be extended in real time when technical difficulties arise. There is real time notification of technical problems to all participants in the exam. This permits fast and efficient resolution of technical problems. In addition, there is real time communication of administrative staff with students. This further enhances the ability to troubleshoot technical problems without destroying the temporal order of the exam (order of stations). One other clear advantage is the reduced need for examination administration staff; five per circuit in our exam setting. This was manageable and we were able to run the exams smoothly. This compares favourably to what is needed in exams delivered by generic video conferencing software ((Shehata et al., 2021)). Thus, we believe that such exams are feasible and even sustainable in the long run and beyond the pandemic as evidence shows that adopting new technological tools in clinical skills teaching and evaluation improves accessibility and efficiency.(O'Rae et al., 2025).

An important limitation of the approach presented here, which, however, applies to all remote exams in general, is the inability to test skills of clinical physical examination. This however, can be partly mitigated by using examination cards and by enhancing relevant WPBA components, such as the Clinical Examination and Procedural Skills (CEPS) component. This has been shown in the case of the membership examination in orthodontics during the COVID-19 pandemic (Eckhardt et al., 2021). Indeed, regardless of the pandemic, WPBAs have now been suggested as a better tool to assess examination skills in the context of real life despite the reliability limitations (Khan, 2017). Thus, the reduced role of SS in examining certain clinical skills can be replaced by well conducted WPBA assessments. A second important limitation of the approach presented here, is the fact that the required number of administrative staff increases with the number of students (and thus additional circuits) that one needs to run. Whether this factor proves to be limiting in larger-scale exams is something that still needs to be established, although as the technology improves over time this limitation is likely to become less severe.

Finally, with respect to the question of whether exams of this sort can be used as a tool even outside the setting of the pandemic, our experience shows that this is indeed the case, at least in our setting of a small number of medical professional examinees.

6. CONCLUSION

The challenges of delivering clinical skills examinations in the setting of the pandemic have led to the development of a purpose-built remote assessment tool that is sufficiently valid and reliable thus allowing for its use to considered outside the setting of the pandemic.

Author Contributions: VI, AH, and OK developed and evaluated assessment delivery process. AS, KH, AH, DV and OK contributed to exam quality assurance. AH and VI analysed psychometric and results data. AH produced the first draft of the manuscript. All authors critically revised the manuscript and approved the final version.

Acknowledgements: The authors wish to acknowledge the support of Dr Thomas Kropmans and the Qpercom team for training on the use of the software and the development and delivery of the assessments.

Funding: This study was not supported by any specific funding.

Competing interests:
The authors acknowledge that they have worked together with Qpercom staff to receive training on the use of the software and to develop and deliver the assessments. However, the authors confirm that they have not received any financial compensation for this work that could have influenced its outcome in any way. The authors alone are responsible for the content and writing of this article.

REFERENCES

Almarzooq, Z. I., Lopes, M., & Kochar, A. (2020). Virtual learning during the COVID-19 pandemic: A disruptive technology in graduate medical education. *Journal of the American College of Cardiology, 75*(20), 2635–2638. https://doi.org/10.1016/j.jacc.2020.04.015

Botan, V., Laparidou, D., Phung, V.-H., Cheung, P., Freeman, A., Wakeford, R., Denney, M., Law, G. R., & Siriwardena, A. N. (2023). Examiner perceptions of the MRCGP recorded consultation assessment for general practice licensing during COVID-19: Cross-sectional study. *BMC Medical Education, 23*(1), 65. https://doi.org/10.1186/s12909-023-04027-4

Chan, S. C. C., Choa, G., Kelly, J., Maru, D., & Rashid, M. A. (2023). Implementation of virtual OSCE in health professions education: A systematic review. *Medical Education, 57*(9), 833–843. https://doi.org/10.1111/medu.15089

Eckhardt, C. E., Seehra, J., Chadwick, S. M., Voerman, K., Landau, A., Ryan, F. S., Fleming, P. S., Garrett, M., & Cobourne, M. T. (2021). Development, implementation and feedback for an online speciality membership examination in orthodontics during the COVID-19 pandemic. *British Dental Journal, 231*(8), 503–511. https://doi.org/10.1038/s41415-021-3535-5

Giri, J., & Stewart, C. (2023). Innovations in assessment in health professions education during the COVID-19 pandemic: A scoping review. *The Clinical Teacher, 20*(5), e13634. https://doi.org/10.1111/tct.13634

Hamamoto Filho, P. T., Cecilio-Fernandes, D., Norcia, L. F., Sandars, J., Anderson, M. B., & Bicudo, A. M. (2022). Reduction in final year medical students' knowledge during the COVID-19 pandemic: Insights from an interinstitutional progress test. *Frontiers in Education, 7.* https://www.frontiersin.org/articles/10.3389/feduc.2022.1033732

Harvill, L. M. (1991). Standard error of measurement. *Educational Measurement: Issues and Practice, 10*(2), 33–41. https://doi.org/10.1111/j.1745-3992.1991.tb00195.x

Hilburg, R., Patel, N., Ambruso, S., Biewald, M. A., & Farouk, S. S. (2020). Medical education during the coronavirus disease-2019 pandemic: Learning from a distance. *Advances in Chronic Kidney Disease, 27*(5), 412–417. https://doi.org/10.1053/j.ackd.2020.05.017

Hopayian, K., Papageorgiou, A., Hadgipapas, A., Charalambous, A., & Kolokotroni, O. (2019). Teaching family medicine online: The experience in Cyprus. *British Journal of General Practice, 69*(682), 244–244. https://doi.org/10.3399/bjgp19X702461

Kane, M. (2009). Validating the interpretations and uses of test scores. In *The concept of validity: Revisions, new directions, and applications* (pp. 39–64). IAP Information Age Publishing.

Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin, 112*(3), 527–535. https://doi.org/10.1037/0033-2909.112.3.527

Khan, H. (2017). OSCEs are outdated: Clinical skills assessment should be centred around workplace-based assessments (WPBAs) to put the 'art' back into medicine. *MedEdPublish, 6,* 189. https://doi.org/10.15694/mep.2017.000189

Kunutsor, S. K., Metcalf, E. P., Westacott, R., Revell, L., & Blythe, A. (2022). Are remote clinical assessments a feasible and acceptable method of assessment? A systematic review. *Medical Teacher, 44*(3), 300–308. https://doi.org/10.1080/0142159X.2021.1987403

Messick, S. (1989). Validity. In *Educational measurement* (3rd ed., pp. 13–103). American Council on Education.

O'Rae, A., Peters, K., Shajani, Z., Burkett, J., & Laing, C. (2025). Improving the evaluation of clinical competence in undergraduate students, evidence and technology: An integrative review. *Journal of Professional Nursing, 56,* 19–25. https://doi.org/10.1016/j.profnurs.2024.10.003

Patrício, M. F., Julião, M., Fareleira, F., & Carneiro, A. V. (2013). Is the OSCE a feasible tool to assess competencies in undergraduate medical education? *Medical Teacher, 35*(6), 503–514. https://doi.org/10.3109/0142159X.2013.774330

Quattrocchi, A., Mamais, I., Tsioutis, C., Christaki, E., Constantinou, C., Koliou, M., Pana, Z.-D., Silvestros, V., Theophanous, F., Haralambous, C., Stylianou, A., Sotiriou, S., Athanasiadou, M., Kyprianou, T., Demetriou, A., Demetriou, C. A., Kolokotroni, O., Gregoriou, I., Paphitou, N., … Nikolopoulos, G. (2020). Extensive testing and public health interventions for the control of COVID-19 in the Republic of Cyprus between March and May 2020. *Journal of Clinical Medicine, 9*(11), Article 11. https://doi.org/10.3390/jcm9113598

Ryan, A., Carson, A., Reid, K., Smallwood, D., & Judd, T. (2020). Fully online OSCEs: A large cohort case study. *MedEdPublish, 9*(1). https://doi.org/10.15694/mep.2020.000214.1

Saad, S. L., Richmond, C., Jones, K., Schlipalius, M., Rienits, H., & Malau-Aduli, B. S. (2022). Virtual OSCE delivery and quality assurance during a pandemic: Implications for the future. *Frontiers in Medicine, 9*. https://www.frontiersin.org/articles/10.3389/fmed.2022.844884

Servin-Rojas, M., Olivas-Martinez, A., Dithurbide-Hernandez, M., Chavez-Vela, J., Petricevich, V. L., García-Juárez, I., Gallo de Moraes, A., & Zendejas, B. (2022). Impact of the COVID-19 pandemic on the clinical training of last year medical students in Mexico: A cross-sectional nationwide study. *BMC Medical Education, 22*(1), 24. https://doi.org/10.1186/s12909-021-03085-w

Shaban, S., Tariq, I., Elzubeir, M., Alsuwaidi, A. R., Basheer, A., & Magzoub, M. (2021). Conducting online OSCEs aided by a novel time management web-based system. *BMC Medical Education, 21*(1), 508. https://doi.org/10.1186/s12909-021-02945-9

Shehata, M. H., Kumar, A. P., Arekat, M. R., Alsenbesy, M., Mohammed Al Ansari, A., Atwa, H., Ahmed, S. A., & Deifalla, A. (2021). A toolbox for conducting an online OSCE. *The Clinical Teacher, 18*(3), 236–242. https://doi.org/10.1111/tct.13285

Trejo-Mejía, J. A., Sánchez-Mendiola, M., Méndez-Ramírez, I., & Martínez-González, A. (2016). Reliability analysis of the objective structured clinical examination using generalizability theory. *Medical Education Online, 21,* 10.3402/meo.v21.31650. https://doi.org/10.3402/meo.v21.31650

Wallace, D., Sturrock, A., & Gishen, F. (2021). "You've got mail!": Clinical and practical skills teaching re-imagined during COVID-19. *Future Healthcare Journal, 8*(1), e50–e53. https://doi.org/10.7861/fhj.2020-0231