

Very Short Answer Questions: A Novel Approach To Summative Assessments In Pathology

This article was published in the following Dove Press journal:
Advances in Medical Education and Practice

Amir H Sam ¹
Emilia Peleva ¹
Chee Yeen Fung¹
Nicki Cohen²
Emyr W Benbow³
Karim Meeran ¹

¹Imperial College School of Medicine, Imperial College London, London, UK;

²King's College London, London, UK;

³University of Manchester, Manchester, UK

Background: A solid understanding of the science underpinning treatment is essential for all doctors. Pathology teaching and assessment are fundamental components of the undergraduate medicine curriculum. Assessment drives learning and the choice of assessments influences students' learning behaviours. The use of multiple-choice questions is common but is associated with significant cueing and may promote "rote learning". Essay-type questions and Objective Structured Clinical Examinations (OSCEs) are resource-intensive in terms of delivery and marking and do not allow adequate sampling of the curriculum. To address these limitations, we used a novel online tool to administer Very Short Answer questions (VSAQs) and evaluated the utility of the VSAQs in an undergraduate summative pathology assessment.

Methods: A group of 285 medical students took the summative assessment, comprising 50 VSAQs, 50 single best answer questions (SBAQs), and 75 extended matching questions (EMQs). The VSAQs were machine-marked against pre-approved responses and subsequently reviewed by a panel of pathologists, with the software remembering all new marking judgements.

Results: The total time taken to mark all 50 VSAQs for all 285 students was 5 hours, compared to 70 hours required to manually mark an equivalent number of questions in a paper-based pathology exam. The median percentage score for the VSAQs test (72%) was significantly lower than that of the SBAQs (80%) and EMQs (84%), $p < 0.0001$. VSAQs had a higher Cronbach alpha (0.86) than SBAQs (0.76), and EMQs (0.77). VSAQs, SBAQs and EMQs had a mean point-biserial of 0.35, 0.30 and 0.28, respectively.

Conclusion: VSAQs are an acceptable, reliable and discriminatory method for assessing pathology, and may enhance students' understanding of how pathology supports clinical decision-making and clinical care by changing learning behaviour.

Keywords: pathology, teaching, assessment, very short answer questions

Introduction

Described as the "science underpinning medicine",¹ pathology is fundamental for all doctors, helping to guide clinical reasoning, the appropriate use and interpretation of laboratory tests, accurate diagnoses and planning patient care.² Information from pathology laboratories is needed for 70% of the diagnoses in hospital inpatients in the United Kingdom (UK).¹ Consequently, pathology teaching should be an integral part of undergraduate medical education. The survey of UK medical schools³ suggests there is great variation in pathology teaching. Some authors have raised concerns about a decrease in pathology teaching in modern medical curricula and its impact on junior doctors' understanding of what is wrong with their patients and their ability to interpret investigation results.²

Correspondence: Karim Meeran
Department of Endocrinology, Charing Cross Hospital, Fulham Palace Road, London W6 8RF, UK
Email k.meeran@imperial.ac.uk

Assessments are known to drive learning.^{4,5} Currently, most undergraduate assessments use multiple-choice questions, such as Single Best Answer questions (SBAQs) or Extended Matching Questions (EMQs),³ whereby candidates are presented with a list of possible answers from which they select the most appropriate response. Well-constructed multiple-choice questions such as SBAQs and EMQs can assess deep learning; however, they have been criticised as these formats test recognition rather than recall⁶ and are subject to cueing.⁷ Furthermore, students will prepare differently for different examination formats.^{8–10} Multiple-choice questions have been shown to elicit test-taking behaviours such as “rote learning”, that may be inauthentic to real-world clinical reasoning;¹¹ patients do not present with a list of five possible diagnoses for the doctor to choose from. If assessments required students to recall knowledge, rather than select responses, this may alter learning behaviour by driving students to seek deeper understanding of the subject. Requiring candidates to generate responses has also been demonstrated to improve long-term retention after studying.^{12–15} Essay-type questions and Objective Structured Clinical Examinations (OSCEs) can test the ability to recall and apply knowledge, and are used by some UK medical schools to assess pathology;³ however, these are very resource-intensive in terms of delivery and marking, and can only cover limited sections of the curriculum.

An alternative assessment method is Very Short Answer questions (VSAQs), consisting of a clinical vignette followed by a question (usually about diagnosis or management), which requires candidates to generate a short response, typically one to four words long.⁷ We have previously shown VSAQs, administered using a novel online assessment management software, to be a highly reliable and discriminatory assessment method in formative examinations; however, these findings need to be confirmed in summative assessments, as the level of motivation of students will be different in high-stakes summative settings.

We used an online tool to run a pathology summative assessment at Imperial College London. The aim of this study was to evaluate whether VSAQs used in a summative pathology assessment were an acceptable, reliable and discriminatory assessment tool.

Methods

Participants And Assessment

The Medical Education Ethics Committee at Imperial College London deemed this study to be an assessment evaluation,

which did not require formal ethical approval. The pathology course at Imperial College School of Medicine currently starts at the beginning of Year 5 (the penultimate year), with a block of teaching, followed by some integrated pathology during Year 5. All medical students in Year 5 ($n=338$ in 2017 and $n=285$ in 2018) undertook a summative pathology assessment, as well as a written paper and a clinical skills assessment focusing on the specialties taught in Year 5. We introduced 25 VSAQs in the Pathology exam in 2017, which were delivered on paper and marked by hand. We subsequently included 50 VSAQs in the Pathology exam in 2018, which were administered on an iPad using an online exam management software (Practique; Fry-IT Ltd, London, UK) along with the Safe Exam Browser software, to ensure that only the exam was visible, and all other websites and applications were disabled.

Questions were written by experienced clinicians familiar with the pathology curriculum. All items were reviewed by external examiners and the standard setting panel. The Ebel method was used to determine the pass mark. The assessment consisted of 175 questions: 50 VSAQs, 50 SBAQs and 75 EMQs. There were 10 VSAQs, 10 SBAQs and 15 EMQs on each topic: haematology, immunology, histopathology, chemical pathology and microbiology. The questions covered the pathology curriculum at Imperial College School of Medicine and were mapped to the Royal College of Pathologists undergraduate medicine curriculum (Table 1).¹ All students answered all questions. The length of the assessment was 180 mins.

Marking

At the end of the online assessment, answers to the VSAQs were machine-marked using a semi-automated algorithm designed to reduce marking time. This compares the student's answer against a pre-defined list of correct answers and uses a measure called the Levenshtein distance to measure how closely a student's given answer matches the pre-approved correct answers. All answers that were identical to the list of approved answers were automatically marked as correct. Students' answers that were within a Levenshtein distance of 3.0, meaning they were within three character errors (insertions, deletions or substitutions) of a preapproved answer, were identified by the algorithm as an approximate match. For example, if a student had answered “thyroditis,” a misspelling of the correct answer “thyroiditis”, this would be identified as an approximate match by the algorithm as it would have a Levenshtein distance of 1.0, since only one character insertion is required to make it an exact match. All answers with a Levenshtein distance of greater than 3.0 were

Table 1 The Pathology Summative Assessment Blueprint Questions Mapped To The Royal College Of Pathologists Undergraduate Curriculum. Each Question Covered One To Three Areas

Knowledge Relating To:	Number Of Questions
Cell biology, pathology and clinical features of biological processes	
Inflammation	34
Infection and its treatment	31
Neoplasia and cancer	31
Circulation	34
Intrauterine and childhood development	8
Occupational, environmental and global health	10
Systems of the body	
Cardiovascular system	8
Respiratory system	10
Gastrointestinal system	23
Endocrine system	19
Central nervous system	6
Renal and genitourinary systems	14
Reproductive systems	3
Skin	9
Musculoskeletal system	21
Blood & immunity systems	48
General competencies: working with pathologists, working with the coroner, working with patients, and clinical governance and patient safety	22
Total number of questions mapped:	175

Note: Data from The Royal College of Pathologists.¹

marked as a non-match by the algorithm. All approximate matches and non-matches were then reviewed by a panel of pathologists (see Figure 1). The panel consisted of four pathologists reviewing all student responses online and agreeing on acceptable answers.

Identical responses were grouped in blocks by the application, and responses marked as correct by the examiners were applied to all identical answers.⁷ Any answers marked as correct by the examiners were automatically added to the set of acceptable responses for that question and the software will recall these responses if the same question is used again in future.

In order to evaluate the acceptability of the VSAQs in terms of faculty time required for marking, we compared the time taken for manual marking of the Pathology paper in 2017 versus the time taken for examiner reviews after online marking in 2018 (time was rounded to the nearest hour). Answers to EMQs and SBAQs were entirely machine-marked.

Analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows Version 24.0 (IBM Corp., Armonk,

NY, USA) and PRISM Version 5.0C (Graphpad Software, Inc., San Diego, CA, USA). Mean is given for normally distributed data, and median for non-normally distributed data. The Kruskal–Wallis test with Dunn’s multiple comparisons test was used for non-normally distributed data to assess the difference between the groups. Cronbach’s alpha was calculated as a measure of reliability. Cronbach’s alpha for EMQs was adjusted for a 50-question test using the Spearman-Brown prediction formula. Item-total score point-biserial was calculated as a measure of discrimination.

Results

Acceptability

Double marking of 25 VSAQs for 338 students in 2017 took 42 hours. The total time spent by examiners to review the machine-marked answers to all 50 VSAQs for 285 students in 2018 was 5 hours.

Facility

In the 2018 exam, the median for the VSAQs was 72% (interquartile range 62%–82%), SBAQs 80% (interquartile

Possible responses:

Acute lymphoblastic leukaemia 1.00

Legend: Correct response Distance below or equal threshold
Distance above threshold No response given

	Response	Number of responses	Score (max. 1.00)	Add as a new answer
Exact match	Acute lymphoblastic leukaemia	92 candidates	1.00	
Match failure	Acute myeloid leukaemia	24 candidates	0.00	
Examiner overrides	Acute lymphocytic leukaemia	21 candidates	1.00	
Non-exact matches	Acute Lymphoblastic Leukaemia	17 candidates	1.00	
	None	10 candidates	0.00	
	Aplastic anaemia	8 candidates	0.00	
	Myelodysplastic syndrome	7 candidates	0.00	

Figure 1 Example of computer-marking of a Very Short Answer question (VSAQ). The answers shown in green match those on the list of possible answers and are automatically assigned a mark (1.00). The amber answer has been marked as correct based on its similarity to the acceptable answer. Answers marked by computer as incorrect are shown in red. During the examiner verification process, this can be over-ridden with all identical answers automatically receiving a mark (e.g., “acute lymphocytic leukaemia” in this case).

range 72%–86%) and EMQs 84% (interquartile range 76%–88%). The median percentage score for the VSAQs test was significantly lower than both SBAQs and EMQs ($p < 0.0001$).

Reliability And Discrimination

Cronbach’s alpha for the VSAQs test was 0.86, compared to 0.76 for the SBAQs test and 0.77 for the EMQs test. The mean item-total score point-biserial for the VSAQs was 0.35, compared to 0.30 for the SBAQs and 0.28 for EMQs.

Discussion

An alternative assessment method may be useful in driving pathology learning in medical students. VSAQs are a method for assessing ability to recall and apply knowledge

as well as enabling broad sampling of the pathology curriculum across a wide variety of topics. We have used VSAQs with an online assessment tool for the first time in a summative pathology assessment. Results show that VSAQs can be an acceptable, reliable and discriminatory method for assessment of pathology.

Students scored significantly lower on the VSAQs test, consistent with our previous findings suggesting that students find this assessment format more difficult.⁷ Unlike SBAQs and EMQs, the VSAQs format requires students to generate, rather than recognise, responses and to demonstrate deeper understanding of pathology. Students have previously agreed that VSAQs were more representative of clinical practice and that using VSAQs in summative

examinations will likely influence learning behaviour and improve preparation for clinical practice.⁷

Use of open-ended questions has previously been limited by the resource-intensive nature of administration and marking by examiners. Furthermore, paper-based delivery of open-ended questions for large cohorts of students is limited by the need for decollation of the mark sheets, marking by hand, the entering marks into spreadsheets, and the introduction of the risk of human error. We have shown how to overcome these limitations using a novel online assessment tool and machine-marking. We were able to mark all 50 VSAQs for 285 students on the same day as the exam.

As identical responses were grouped in blocks by the application, responses marked as correct by the examiners were applied to all identical responses. This facilitated the review process, reduced marking time and ensured consistency. Furthermore, the online assessment software remembers new marking judgments and saves these to the existing set of acceptable responses for that question. This ensures that marking time improves with future use of each question.⁷

This study is limited by the sample size and inclusion of students from a single centre only. As more undergraduate programs include VSAQs in their assessments, the utility of this assessment instrument for larger cohorts could be evaluated. Another limitation is that student feedback was not collected for this study; however, we have previously reported positive student feedback for VSAQs in a formative exam.⁷

In summary, the use of VSAQs in summative pathology assessments is reliable, discriminatory and acceptable, and the assessment method will likely encourage deeper learning of pathology by undergraduate students. Future studies need to explore the impact of VSAQs on deep learning. Choice of assessments can complement teaching strategies, including better signposting to students and making pathology more visibly clinically relevant, enhancing students' engagement with the specialty. Greater engagement with the specialty would enhance students'

understanding of the role of pathology in supporting clinical decision-making and clinical care.

Disclosure

The authors report no conflicts of interest in this work.

References

1. The Royal College of Pathologists. Pathology Undergraduate Curriculum. 2014.
2. Marsdin E, Biswas S. Are we learning enough pathology in medical school to prepare us for postgraduate training and examinations? *J Biomed Educ.* 2013. doi:10.1155/2013/165691
3. Mattick K, Marshall R, Bligh J. Tissue pathology in undergraduate medical education: atrophy or evolution? *J Pathol.* 2004;203(4):871–876. doi:10.1002/path.v203:4
4. Cox M, Irby DM, Epstein RM. Assessment in medical education. *N Engl J Med.* 2007;356(4):387–396. doi:10.1056/NEJMra054784
5. Wass V, Van Der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. *Lancet.* 2001;357:945–949. doi:10.1016/S0140-6736(00)04221-5
6. Veloski JJ, Rabinowitz HK, Robeson MR, Young PR. Patients don't present with five choices: an alternative to multiple-choice tests in assessing physicians' competence. *Acad Med.* 1999;74(5):539–546. doi:10.1097/00001888-199905000-00022
7. Sam AH, Field SM, Collares CF, et al. Very-short-answer questions: reliability, discrimination and acceptability. *Med Educ.* 2018;52(4):447–455. doi:10.1111/medu.13504
8. Cilliers FJ, Schuwirth LWT, Van Der Vleuten CPM. A model of the pre-assessment learning effects of assessment is operational in an undergraduate clinical context. *BMC Med Educ.* 2012;12:9. doi:10.1186/1472-6920-12-9
9. Al-Kadri HM, Al-Moamary MS, Roberts C, Van Der Vleuten CPM. Exploring assessment factors contributing to students' study strategies: literature review. *Med Teach.* 2012;34:S42–S50. doi:10.3109/0142159X.2012.656756
10. Newble DI, Jaeger K. The effect of assessments and examinations on the learning of medical students. *Med Educ.* 1983;17(3):165–171. doi:10.1111/medu.1983.17.issue-3
11. Surry LT, Torre D, Durning SJ. Exploring examinee behaviours as validity evidence for multiple-choice question examinations. *Med Educ.* 2017. doi:10.1111/medu.13367
12. McConnell MM, St-Onge C, Young ME. The benefits of testing for learning on later performance. *Adv Health Sci Educ.* 2014;20(2):305–320. doi:10.1007/s10459-014-9529-1
13. Larsen DP, Butler AC, Roediger HL. Test-enhanced learning in medical education. *Med Educ.* 2008;42:959–966. doi:10.1111/med.2008.42.issue-10
14. Wood T. Assessment not only drives learning, it may also help learning. *Med Educ.* 2009;43:5–6. doi:10.1111/med.2008.43.issue-1
15. McDaniel MA, Roediger HL, McDermott KB. Generalizing test-enhanced learning from the laboratory to the classroom. *Psychon Bull Rev.* 2007;200–206. doi:10.3758/BF03194052

Advances in Medical Education and Practice

Dovepress

Publish your work in this journal

Advances in Medical Education and Practice is an international, peer-reviewed, open access journal that aims to present and publish research on Medical Education covering medical, dental, nursing and allied health care professional education. The journal covers undergraduate education, postgraduate training and continuing medical education

including emerging trends and innovative models linking education, research, and health care services. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/advances-in-medical-education-and-practice-journal>