

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Using prescribing very short answer questions to identify sources of medication errors: a prospective study in two UK medical schools
AUTHORS	Sam, Amir; Fung, Chee Yeen; Wilson, Rebecca; Peleva, Emilia; Kluth, David; Lupton, Martin; Owen, David; Melville, Colin; Meeran, Karim

VERSION 1 - REVIEW

REVIEWER	Phil Smith University Hospital of Wales, Cardiff
REVIEW RETURNED	29-Mar-2019

GENERAL COMMENTS	<p>This study addresses an important question: important from the perspective that very short answer questions offer a better assessment of knowledge than single best answer questions, and also important from the perspective of better assessing knowledge of prescribing safety. The study has identified some potentially concerning results, especially in medical students' ability to prescribe some groups of medications correctly (specifically insulins and analgesia).</p> <p>The authors have addressed the main potential weaknesses of the paper, notably that the cohort of students participating in this study may not be representative of the medical student body as a whole. The study points the way forward for improving knowledge assessment in this important area of clinical practice.</p>
-------------------------	---

REVIEWER	Dr Anne Holbrook McMaster University Canada
REVIEW RETURNED	30-Mar-2019

GENERAL COMMENTS	<p>Since medical education aims to ensure competence to perform 'entrustable professional activities', a major one of which is to prescribe effectively, safely and accurately, this is an interesting study.</p> <p>The description of what was done is good and the writing is clear, but requires clarification and additional detail in several areas, I think:</p> <p>a) the question and the intervention are repeatedly referred to as 'prescribing VSA' but in fact the main question is not just about evaluating a 'very short answer' exam intervention, it is equally about evaluating whether a machine-scoring system works or not. This needs to be clarified.</p>
-------------------------	---

	<p>b) More background on short answer questions and very short answer questions - what is the difference, what is the evidence outside of the realm of prescribing that these formats are superior to multiple choice questions, should be added.</p> <p>c) More information on the PRACTIQUE software should be included - primarily how is the computer able to mark a prescription. is this by optical character recognition, etc? Presumably answers are typed (this should be specified). Are answers entered in test fields where only one word is required (eg, the prescription is set up as an electronic fill-in-the-blank template)? How does this software differ from the PSA, which also uses software that can mark a prescription?</p> <p>d) Clarify whether only one correct drug/dose/route,etc was allowed. if this is the case, then this is a limitation, as there is almost never one and only one correct prescription for a condition. The PSA uses a preferable system where some responses will receive full marks, some partial marks, some none at all.</p> <p>e) More description of the clinical scenario, with an example of a scenario and the prescribing template, would be helpful to reassure readers that 'best practices' of scenario writing have been followed.</p> <p>f) The study seems overly focused on the UK environment, background work and intent. It may well be that the UK is the most advanced of countries around evaluating prescribing competencies, however medical education is highly advanced in many other countries. The background should discuss a worldwide literature review on the topic, if not a systematic review. One of the key parts of any discussion section should be explaining how this study adds to the medical literature. this is not possible without a comprehensive literature review.</p> <p>g) The limitations are not fully explained and the resulting conclusion goes beyond the reach of the study. The sample size is magnified by looking at individual prescriptions but prescriptions are correlated across students, and likely even across schools (this should be tested) such that the actual sample size could be the number of students or, even, the number of schools (2). The authors have not adequately explained the limited generalizability of the work. This is essentially a pilot study (an encouraging one with obvious import for other schools) that needs to be replicated much more widely.</p> <p>h) The concept of 'acceptability, reliability, and discrimination' as outlined in the aims are not defined in the methods. It would be helpful to define these and then specifically address these concepts by name in the results.</p>
--	---

REVIEWER	Nagaswami Vasam Cooper Medical School of Rowan University, USA
REVIEW RETURNED	28-Apr-2019

GENERAL COMMENTS	<p>Using prescribing very short answer questions to identify sources of medication errors</p> <p>This is well written paper on an important topic of prescribing medications where senior medical students and junior doctors potentially could make errors. The VSA is an excellent way of teaching students how to write a script. Unfortunately, EMR diminished the opportunity for students to practice writing prescriptions.</p>
-------------------------	--

	<p>The introduction is well written explaining the reasons for this study. The authors reviewed and included 29 relevant references.</p> <p>I have the following comments and questions in the methods and results.</p> <ol style="list-style-type: none"> 1. Was there a practice test so that students are familiar with the VSA approach? 2. The students took the VSA first and then the SBA. 3. Since the 'n' of 364 is large enough, I wonder how the outcomes would have been if some of the students were given SBA followed by VSA? Looking at the results, students had difficulties in 5 areas out of 10. 4. It is in the best interest of the readers of the article I strongly suggest that the authors consider providing 2-3 examples of the case scenario. <p>Discussion: Page 10, para 1, lines 15-16: "Additionally, personalized feedback can be sent out to students" ... : I wonder whether the student actually received such feedback?</p> <p>In all, this is an excellent approach to teach students about correct way to prescribe. Unfortunately, as the authors mentioned because of the electronic approach teaching important things cannot be done to the faculty satisfaction.</p>
--	---

VERSION 1 – AUTHOR RESPONSE

REVIEWER 1

We thank reviewer 1 for their comments and are extremely grateful for his positive review of the manuscript.

REVIEWER 2

We thank reviewer 2 for their helpful review of our manuscript. We have addressed the points raised as detailed below.

1. The question and the intervention are repeatedly referred to as 'prescribing VSA' but in fact the main question is not just about evaluating a 'very short answer' exam intervention, it is equally about evaluating whether a machine-scoring system works or not. This needs to be clarified.

Response: We apologise that this was not clear in the original manuscript. We have revised the stated aims of the study to explicitly state that as well as evaluating the VSA question format, we are also evaluating the machine-scoring system for marking VSAs on a large scale (please see page 2, lines 29-30 and page 6, lines 155-156).

2. More background on short answer questions and very short answer questions - what is the difference, what is the evidence outside of the realm of prescribing that these formats are superior to multiple choice questions, should be added.

Response: We apologise for not including sufficient detail in the original manuscript. We have expanded on the difference between SAQs and VSAs (please see page 6, lines 141-7). We have added the evidence that VSAs are potentially more reliable and discriminatory than multiple choice questions when used outside of prescribing (please see page 6, lines 140-1).

3. More information on the PRACTIQUE software should be included - primarily how is the computer able to mark a prescription. Is this by optical character recognition, etc? Presumably answers are typed (this should be specified). Are answers entered in test fields where only one word is required (eg, the prescription is set up as an electronic fill-in-the-blank template)? How does this software differ from the PSA, which also uses software that can mark a prescription?

Response: We apologise for not including this information in the original manuscript. We have added further detail as to how PRACTIQUE marks the prescription (please see page 8, lines 192-6) and have clarified exactly how the students were required to enter their answers (please see page 7, p174-6). Our system slightly differs from the PSA software which enables students to select the drug name from a dropdown menu.

4. Clarify whether only one correct drug/dose/route,etc was allowed. If this is the case, then this is a limitation, as there is almost never one and only one correct prescription for a condition. The PSA uses a preferable system where some responses will receive full marks, some partial marks, some none at all.

Response: We apologise that this was not clear in the original manuscript. We have clarified that more than one drug/dose/route deemed acceptable by the examiners was included in the list of preapproved answers (please see page 8, lines 196-7).

5. More description of the clinical scenario, with an example of a scenario and the prescribing template, would be helpful to reassure readers that 'best practices' of scenario writing have been followed.

Response: We are grateful to the reviewer for highlighting this issue – we have now included two example clinical scenarios in the supplementary material.

6. The study seems overly focused on the UK environment, background work and intent. It may well be that the UK is the most advanced of countries around evaluating prescribing competencies, however medical education is highly advanced in many other countries. The background should discuss a worldwide literature review on the topic, if not a systematic review. One of the key parts of any discussion section should be explaining how this study adds to the medical literature. this is not possible without a comprehensive literature review.

Response: We thank the reviewer for highlighting this omission and apologise for not including this information in the original manuscript. We have now referred to the worldwide issue of assessing prescribing competency in medical students and included additional references to a large European cross-sectional study and an international systematic review (please see p5, lines 106-7) We have also highlighted that the PSA is being adopted in countries outside of the U.K. (please see p5, line 119-20).

7. The limitations are not fully explained and the resulting conclusion goes beyond the reach of the study. The sample size is magnified by looking at individual prescriptions but prescriptions are correlated across students, and likely even across schools (this should be tested) such that the actual sample size could be the number of students or, even, the number of schools (2). The authors have not adequately explained the limited generalizability of the work. This is essentially a pilot study (an encouraging one with obvious import for other schools) that needs to be replicated much more widely.

Response: We thank the reviewer for highlighting this pertinent point. We have revised the limitations section accordingly. We clarified that this is a pilot study with limited generalisability and further work with larger sample sizes is necessary. We have now emphasised that weaker students are likely to make the same category of error repeatedly, and students within the same institution may make the same error due to a curricula or teaching issue, which may in turn distort the results and limit the generalisability of our findings (please see p12, lines 303-8). We have also removed the reference to the number of prescriptions in the 'Strengths and Limitations' section of the paper.

8. The concept of 'acceptability, reliability, and discrimination' as outlined in the aims are not defined in the methods. It would be helpful to define these and then specifically address these concepts by name in the results.

Response: We thank the reviewer for highlighting this omission and have revised the manuscript accordingly. We have defined acceptability (please see p8, lines 203-4) and addressed this concept in the results (please see p9, lines 220-1). The statistical methods to assess reliability have been described (please see p8, lines 212-3) and explicitly referred to in the results (please see p9, lines 225-6). We have removed the reference to discrimination in the aims.

REVIEWER 3

We thank reviewer 3 for their comments and positive review of the manuscript. We have addressed the points raised as detailed below:

1. Was there a practice test so that students are familiar with the VSA approach?

Response: We thank the reviewer for highlighting this important omission. We have now clarified that all students had previously been exposed to the VSA question format through use of these questions in formative assessments (please see p7, lines 169-70).

2. The students took the VSA first and then the SBA. Since the 'n' of 364 is large enough, I wonder how the outcomes would have been if some of the students were given SBA followed by VSA? Looking at the results, students had difficulties in 5 areas out of 10.

Response: We thank for the reviewer for raising this pertinent point. In previous studies comparing VSA with SBA questions, albeit outside of realm of prescribing (please see reference 30), there has been a significant positive cueing effect associated with sitting the SBA first followed by the VSA. This was the reason we decided that all students should sit the VSA before the SBA.

3. It is in the best interest of the readers of the article I strongly suggest that the authors consider providing 2-3 examples of the case scenario.

Response: We apologise for not including these in the original submission; example scenarios are now provided in the supplementary material.

4. Discussion: Page 10, para 1, lines 15-16: "Additionally, personalized feedback can be sent out to students" ... : I wonder whether the student actually received such feedback?

Response: Unfortunately, we were unable to provide the study participants with individual feedback as the software functionality was not complete at the time; however, this feature has now been incorporated into Practique and in future could be utilised.

VERSION 2 – REVIEW

REVIEWER	Ddr Anne Holbrook,MD,PharmD,MSc,FRCP McMaster University Canada
REVIEW RETURNED	02-Jun-2019

GENERAL COMMENTS	My previous comments were adequately addressed, except one. I still do not see evidence of a comprehensive background international literature review on the use of the very short answer format in medical exams with or without machine readability. this would be a useful addition.
-------------------------	---

REVIEWER	Nagaswami Vasan Cooper Medical School of Rowan University USA
REVIEW RETURNED	29-May-2019

GENERAL COMMENTS	Thank you for responding to my comments (reviewer 3) and also the detailed response to reviewer 2. I am satisfied with the responses to comments. Hence I recommend acceptance of the manuscript.
-------------------------	--

VERSION 2 – AUTHOR RESPONSE

Response to reviewers 2

Many thanks to the reviewers for their helpful comments. We apologise for not clarifying this point in the revisions. Very Short Answers questions (VSAs) is a term coined by our group here at Imperial College London. Therefore, to the best of our knowledge, our two previous papers are the only published literature regarding VSAs. Both these papers are cited in the manuscript. A large study comparing the VSA and SBA question formats across a number of medical schools has recently been conducted, but the results of this are yet to be published.