Health Professions Education

Volume 8 | Issue 2 Article 1

2022

The Hidden Value of Multisource Progress Assessment in Medical Education

Jerome I. Rotgans
Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore. Institution for Medical Education Research Rotterdam, Erasmus University Medical Center, The Netherlands

Muhammad Raihan Jumat
Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore,
raihanjumat@gmail.com

Henk G. Schmidt
Institution for Medical Education Research Rotterdam, Erasmus University Medical Center, The Netherlands

EDITORIAL

The Hidden Value of Multisource Progress Assessment in Medical Education

Jerome I. Rotgans [b], Muhammad R. Jumat [a,*], Henk G. Schmidt [b]

[a] Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore
[b] Institute of Medical Education Research Rotterdam, Erasmus University Medical Center, the Netherlands

Abstract

The COVID-19 pandemic constituted a stress test for medical education, especially for assessment. Exams needed to be postponed or even cancelled. This was a problem because many schools rely on decision making with regard to student performance and their progress through the curriculum using only a few high-stake examinations. In this report we present how a medical school can make a relatively easy transition to multisource progress assessment, based on results on a large number of low-stake tests. Such tests are often integral parts of the curriculum, but hidden because they are not used for decision making. We will present how results of such low-stake tests can be combined and visualized to allow for more comprehensive (and potentially more reliable and valid) assessment decisions.

Keywords: Multisource progress assessment, Medical education, Assessment strategy, Student feedback, COVID-19

1. Background

The significant disruptions the COVID-19 pandemic caused to medical education revealed hidden flaws in the curriculum that would otherwise have gone unnoticed. This is particularly true for assessment. Many medical curricula rely on a limited number of high-stake examinations, such as end-of-year examinations or high-stake OSCEs that must be passed before students can proceed. The pandemic has taught us that relying on a limited number of high-stake examinations can be problematic. Due to lock-downs and safe-distancing measures, students were often not able to physically take examinations, and as a consequence, exams were postponed or even scrapped. Needless to say, having to rely on a limited number of assessments does not allow for making valid decisions about student performance and progress. This is not only true in times of pandemics but applies to decision making in general. The general rule is that the more data are available the more valid decisions are [1]. This is even the case when individual data are less reliable. This point of view has recently been extensively promoted by Van der Vleuten and Schuwirth [2] by their programmatic assessment proposal. They argue that programmatic assessment is based on continual collection of assessment and feedback information which is captured in a portfolio and periodically discussed with a mentor or coach. In this approach, students reflect on their performance and analyze their strengths and weaknesses in different competency domains, rather than purely relying on numerical assessment data (p.2) [3].

In this report we will illustrate how a medical school can transfer its assessment program from being dependent on a limited number of high-stake examinations to multisource progress assessment, relying on many low-stake tests to paint a more comprehensive picture of students' performance and make informed decisions about progress. Our message is that this transition is not difficult because many medical schools have access to a wealth of low-stake tests that tend to be hidden in the curriculum.

Received 27 April 2022; revised 13 July 2022; accepted 20 July 2022.
Available online 15 November 2022

* Corresponding author at: Lee Kong Chian School of Medicine, Nanyang Technological University, Clinical Sciences Building, 11 Mandalay Road, 308232, Singapore.
E-mail address: raihanjumat@gmail.com (M.R. Jumat).

https://doi.org/10.55890/2452-3011.1019

In this report we will demonstrate how to unearth these hidden data and use them for transitioning to more comprehensive assessment.

2. Three drivers for multisource progress assessment

As the name suggests, multisource progress assessment (MPA) makes use of all available sources of assessment, such as low-stake in-course assignments, anatomy spot-tests, quizzes, structured patient encounters, project reports and high-stake examinations. In fact, all assessments are included—no matter how small—to provide a comprehensive overview of how students are performing in almost real time (in the section that follows, we will highlight how these data can be summarized and presented). Having access to a comprehensive track record allows for early intervention and objective decision making about students' progress.

Although the idea of multisource progress assessment is not new, few medical schools have embraced and implemented it [4]. In our view, the COVID-19 pandemic was a wake-up call that made it clear that a change is needed from an overreliance on infrequent high-stake summative examinations to the inclusion of frequent low-stake tests [5]. The realization that change is needed occurred first when students were not allowed to participate physically in examinations due to safe-distancing measures and alternative solutions needed to be found to assess students. Most medical schools resorted to one of two solutions. A low-tech solution was to change the previously closed-book examinations to open-book examinations [6]. However, this also changed the nature of the test since students could now make use of notes and books during the exam. A more technical solution was the adaptation of online proctoring software to assure that students could not cheat when they sat the online examination [7].

At this point it should be noted that high-stake examinations—online or conventional—are a significant source of stress and are the major cause of anxiety in medical students [8]. Changing from high-stake to frequent low-stake tests can address this issue and is therefore an important argument for embracing multisource progress assessment.

Finally, there is a third reason to change to frequent low-stake tests. There is often a misconception regarding the psychometric superiority of high-stake examinations [9]. It is wrongly assumed that high-stake examinations are more reliable than low-stake tests. It is true that high-stake examinations need to be highly reliable (i.e., the results need to be accurate and reproducible) because its consequences for students are so far-reaching. On the other hand, the reliability of low-stake tests are considered less critical. However, if results of many low-stake tests are *combined*, the overall number of items included in decisions is much higher and therefore more valid and reliable [2].

In light of the above drivers, we propose that a transition from over-reliance on a few high-stake examinations to frequent low-stake tests is due. Such approach allows the school to provide timely feedback to the learners and measure their progress more objectively and comprehensively. How this can be accomplished will be exemplified in the next section.

3. Case study: multisource progress assessment

Suppose a medical school would like to implement multisource progress assessment in its curriculum. The curriculum is a five-year MBBS program and after implementation of the first year, multisource progress assessment will be rolled out for the subsequent years—one new cohort at a time. The project is structured along four implementation steps. We will describe these steps in chronological order and provide an overview of the decisions and actions each step requires.

3.1. Step 1: inventory of all assessments

As a first step in the transition process, an inventory has to be made of all assessments for each year. To that end, a simple table can be tabulated by the year heads, capturing the tests, the format, how the tests are scored, frequency of administration etc. See Fig. 1 for a sample.

After the inventory is completed, it is common that there are 400-500 low-stake tests over a five-year MBBS program. All tests within a year need to be completed satisfactorily and passing them is a prerequisite for admission to the end-of-year examinations (i.e., written exam and OSCE).

3.2. Step 2: mapping all tests on competency domains

After the test inventory is completed, the tests are mapped onto the core competency framework of the medical school. In this example, the core competencies consist of five main categories, referred to as competency domains: (1) medical knowledge; (2) clinical/practical skills; (3) communication skills; (4) professionalism; and (5) PBL (problem-based

Year 1 example: Competency map with test inventory

Competency domain	Test	Frequency	Format	Scores
Medical knowledge	End of module progress test	4	True/false questions	Percentage score
	Anatomy spot tests	5	Multiple-choice questions	Percentage score
	Weekly quizzes	40	Multiple-choice questions	Percentage score
	End of year exam	1	Multiple-choice questions	Percentage score
Practical skills	Group observed structured encounter	3	3-point scoring system	Converted percentage score
	Mini clinical evaluation exercise	2	3-point scoring system	Converted percentage score
	Clinical methods test	1	3-point scoring system	Converted percentage score
	Clinical procedures test	1	3-point scoring system	Converted percentage score
	Direct observation of procedural skills test	1	3-point scoring system	Converted percentage score
	End of year OSCE	1	3-point scoring system	Converted percentage score
Communication skills	Patient encounter	2	3-point scoring system	Converted percentage score
	History tacking	2	3-point scoring system	Converted percentage score
	Reflective writing assignment	2	5-point rubric scoring system	Converted percentage score
	Multisource feedback of peers	4	7-point Likert scale	Converted percentage score
Professionalism	Self-evaluation	5	7-point Likert scale	Converted percentage score
	Peer-evaluation	5	7-point Likert scale	Converted percentage score
	Absence score	1	Absence score	Percentage score
	Cooptative science project	2	3-point scoring system	Converted percentage score
	Portfolio	6	3-point scoring system	Converted percentage score
Problem-based learning	Weekly TBL quizzes	48	Multiple-choice questions	Percentage score

Fig. 1. Test inventory year 1.

learning) for the first two years. Tests are then allocated under each competency domain.

Allocating all tests to competency domains is done for two reasons. First, data reduction. Presenting all tests to students is confusing and can lead to data overload. Grouping tests under five competency domains provides structure and students can immediately see how well they are performing for each competency domain since test scores within a competency domain are aggregated to form a composite test score. Second, the same competency domains are used in many post-graduate programs. We believe that familiarizing students with these competency domains early in the undergraduate program will make the transition easier between undergraduate to post-graduate medical training.

3.3. Step 3: developing an assessment dashboard

Once the tests are mapped onto the competency domains, we generate an assessment dashboard that visually displays all test information per competency domain for each student separately. For an overview of the assessment dashboard see Fig. 2. Please take note that the assessment dashboard is only a mock up based on fictitious assessment data to simulate how it will look like.

On the left-hand side, a competency graph is presented to the student, displaying (in this case) his/her performance relative to the cohort's performance (in the form of box plots). A color indication is given to signify if the student is performing well (green) and where he/she stands relative to the cohort. If the student is not performing well (lower quartile) a red marker is provided. However, only presenting aggregated test performance per competency domain can hide important information since the student does not know on which test(s) he/she is not performing well (for this particular student "professionalism"). To provide more information, a detailed performance breakdown is provided in the adjacent window, next to the competency graph on the right. In this window all tests are summed up and populated as the students progress through the year. Finally, also qualitative feedback is provided in the right window for each competency domain. We believe that qualitative feedback is essential for helping students progress and will increasingly be part of all tests (i.e., examiners are required to provide comprehensive qualitative feedback for all tests) [10, 11].

Besides the overview of students' academic performance, many medical school tracks students' professionalism (e.g., being on time for classes). This part of professionalism can be portrayed by a traffic light display that can also be included in the dashboard. Lastly, we propose to include easy access to support staff, such as tutors, study skills counselors and the year heads under the "Need help?" section of the assessment dashboard.

3.4. Step 4: decision-making regarding performance and progress

The assessment dashboard provides a comprehensive overview of students' performance. This

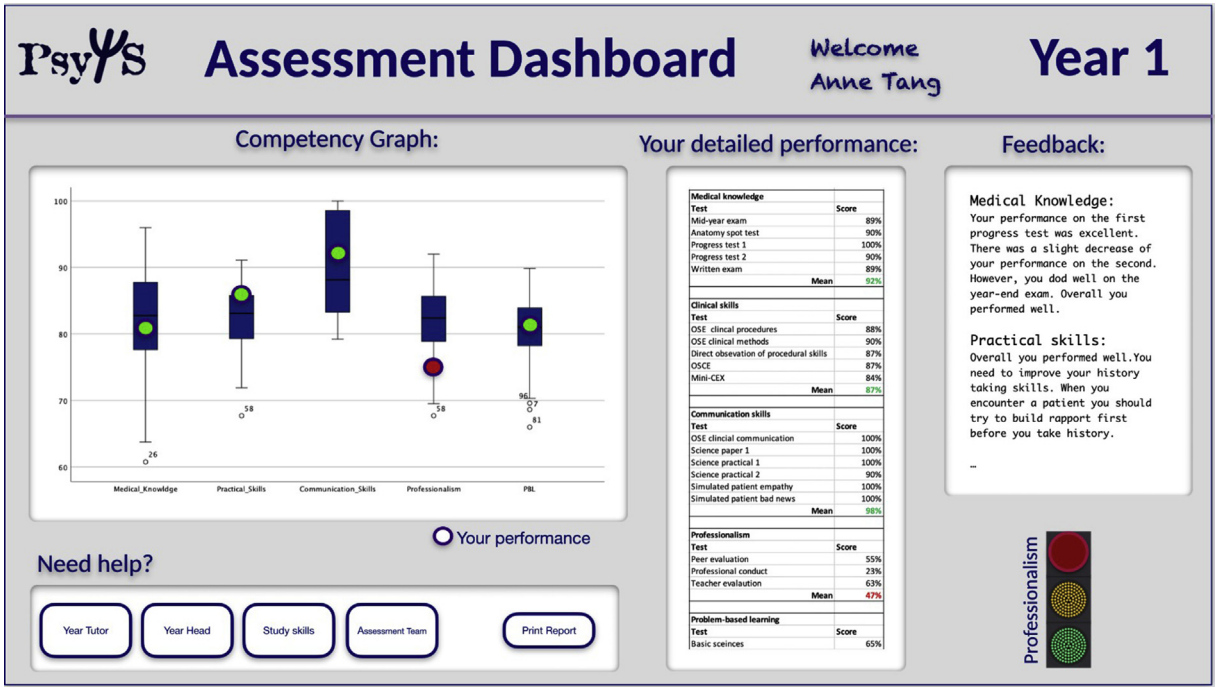

Fig. 2. Mock-up assessment dashboard year 1. Note: the information displayed is from a fictitious student.

information is primarily of importance for the students but also for faculty to track their progress and take action if necessary. The year heads, course leads and instructors have access to the dashboard to examine whether students are struggling. Since the information is updated almost in real time, the dashboard is a useful tool to track progress and intervene in a timely manner. At the end of each module, a formal assessment meeting is convened during which each student's performance is evaluated by the course leads and the faculty. Students will not fail a term, but may receive addition learning opportunities to eliminate any deficiencies. Finally, a formal year-end assessment meeting will be held to decide on students' progress to the next year. During the early implementation of MPA, this final decision will be based on the current assessment rules and regulations, i.e., no changes need to be made at this point, students will have to pass all low-stake tests and high-stake exams. This makes the implementation of MPA administratively easier and allows for making gradual changes to the rules and regulations as MPA progresses (e.g., in the future students may be allowed to compensate on low-stake tests).

Since all information is at the committee's fingertips, a more comprehensive evaluation can be conducted of each student during and at the end of the year; nothing really comes as a surprise since a historical track record is available that documents students' detailed performance. Making decisions in such a holistic manner is based on richer information than relying solely on a few cut-off scores of a limited number of examinations at the end of the year, when it is too late for any improvement action. Final judgments are based on a wealth of information, collected over an entire year, adequately substantiating the decision taken for each student.

4. Conclusions

In this report we presented a straightforward approach of how a medical school can take steps to wean off from an over-reliance on high-stake examinations. The pandemic made us aware of the fact that having to depend on a limited number of high-stake examinations is problematic when one intends to make informed decisions about students' performance and progress. The good news is that many medical schools have access to low-stake tests that are somewhat hidden in the curriculum. In this report we described how results on these tests can be combined and displayed to provide a more complete picture of students' performance enabling making more holistic decisions about their progress. Accessing and organizing these data turned out to be easy since there are many data visualization programs available.

Ethical approval

Ethical approval is not required for this manuscript.

Funding

None.

Conflict of interest

The authors declare no competing interests.

References

[1] Fitzpatrick AR, Yen WM. The effects of test length and sample size on the reliability and equating of tests composed of constructed-response items. Appl Meas Educ 2001;14(1):31–57.
[2] Schuwirth LW, Van Der Vleuten CP. Current assessment in medical education: programmatic assessment. J Appl Testing Technol 2019;20(S2):2–10.
[3] Schut S, Maggio LA, Heeneman S, Van Tartwijk J, Van Der Vleuten C, Driessen E. Where the rubber meets the road — an integrative review of programmatic assessment in health care professions education. Perspect Med Educ 2021;10(1):6–13. https://doi.org/10.1007/s40037-020-00625-w.
[4] van der Vleuten CP, Schuwirth L, Driessen E, Dijkstra J, Tigelaar D, Baartman L, et al. A model for programmatic assessment fit for purpose. Med Teach 2012;34(3):205–14.
[5] Bala L, van der Vleuten C, Freeman A, Torre D, Heeneman S, Sam AH. COVID-19 and programmatic assessment. Clin Teach 2020;17(4):420–2.
[6] Eurboonyanun C, Wittayapairoch J, Aphinives P, Petrusa E, Gee DW, Phitayakorn R. Adaptation to open-book online examination during the COVID-19 pandemic. J Surg Educ 2021;78(3):737–9.
[7] Andreou V, Peters S, Eggermont J, Wens J, Schoenmakers B. Remote versus on-site proctored exam: comparing student results in a cross-sectional study. BMC Med Educ 2021;21(1):1–9.
[8] Aziz N, Serafi AH. Increasing levels of test anxiety and psychological distress with advancing years of medical education. Br J Med Health Res 2017;4(3):40–2.
[9] Schuwirth L, Van der Vleuten C, Durning S. What programmatic assessment in medical education can learn from healthcare. Perspect Med Educ 2017;6(4):211–5.
[10] Bakke BM, Sheu L, Hauer KE. Fostering a feedback mindset: a qualitative exploration of medical students' feedback experiences with longitudinal coaches. Acad Med 2020;95(7):1057–65.
[11] Harrison CJ, Könings KD, Dannefer EF, Schuwirth LWT, Wass V, Van Der Vleuten CPM. Factors influencing students' receptivity to formative feedback emerging from different assessment cultures. Perspect Med Educ 2016;5(5):276–84. https://doi.org/10.1007/s40037-016-0297-x.