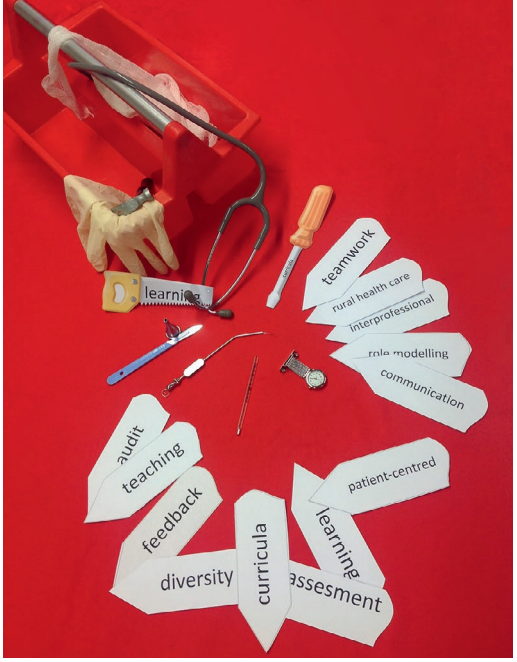
The Clinical Teacher's Toolbox




# Creating a learner performance dashboard for programmatic assessment

**Sally A Santen[1], Leif Myklebust[2], Clare Cabrera[2] Johmarx Patton[2], Mark Grichanik[3], and Nikki L Bibler Zaidi[2]**

[1]Department of Emergency Medicine, School of Medicine, Virginia Commonwealth University, Richmond, Virginia, USA
[2]University of Michigan Medical School, Ann Arbor, Michigan, USA
[3]Rush Medical College of Rush University, Chicago, Illinois, USA

*Editor's note: Programmatic assessment is a process that is gaining traction in medical and health professional education. It is an approach that aligns with competency-based education. Routine and longitudinal data are collected about learners' competence, and their progress is regularly reviewed and analysed. Typically, a variety of assessment methods are used with the overall aim of allowing both learners and educators to gain an understanding of the extent of learning and to maximise feedback for optimal educational impact (assessment for learning). Eventually the process may be used for high-stakes decisions at the end of a programme or phase of study (assessment of learning). In this toolbox article the authors describe how they have developed a method of visually representing medical student attainment during programmatic assessment on a dashboard. This development required input from educators, faculty members and information technology staff. The graphs obtained can help students and staff to see how a student is performing over time. The authors share feedback on the approach from learners and faculty members and provide recommendations for readers who may wish to develop something similar.*

**... we were particularly interested in how data visualisation could help create meaning for stakeholders**

## INTRODUCTION

Medical education is moving towards competency-based medical education (CBME), which must be supported by frequent assessment. The foremost challenge of implementing CBME is developing rigorous assessments with validity evidence that demonstrate a student's progress towards competency.[1,2]

## PROGRAMMATIC ASSESSMENT IN COMPETENCY-BASED MEDICAL EDUCATION

In CBME, assessment data can serve multiple purposes, including the assessment *of* learning (summative competency or grading judgements) and the assessment *for* learning (formative feedback for students).[3] These involve the intentional use of varied assessment methods, resulting in multiple assessments purposefully chosen for their alignment with institutional competencies.[3] Together, these assessment data provide a clearer picture of competent performance.

As programmatic assessment incorporates multiple data points over time, aggregating and displaying these data to provide clear communication for the purpose of feedback on learning and to inform decision making becomes critical. The purpose of this teaching toolbox article is to assist medical and other health professional educators in designing data visualisation for reporting assessment data within a CBME programme. The lessons learned along the journey outlined here will enable educators and administrators to plan, design and implement a data visualisation tool – such as a dashboard – to abstract, summarise or categorise assessment data.[4]

Competency-based medical education assessment can create tension amongst stakeholders, especially students, who hold traditional expectations of norm-referenced grades.[1,2] For example, in the traditional pass/fail grading system of the pre-clerkship phase, a student's performance is generally compared with that of their peers, and cut-off points for passing are commonly determined by class means and standard deviations – this system does not represent CBME assessment. Similarly, there is dissonance between grading and CBME assessment in the clerkships.[2] Therefore, it is important to recognise this tension and to affirm the purpose of the assessment framework.

### Assessment reporting

In programmatic assessment, CBME assessment data must be presented clearly and succinctly to provide meaning for students, faculty members and those responsible for making judgements about competency and progression.[4–6] One challenge involves communicating performance metrics from both summative and formative assessments. Given the importance of displaying and reporting assessment data to stakeholders in a meaningful way, medical schools are partnering with experts in technology and learning analytics to design methods for displaying students' assessments.[7] Data visualisation involves the creation and study of data that have been abstracted, summarised or categorised in some schematic form.[7,8] It combines data from various assessments to organise and visualise data for interpretation, and is employed to effectively provide feedback to students and to assist faculty members in making competency-based decisions.[4,6] The data visualisation can provide efficient feedback, yet Hauer and colleagues found that students' use of this information to guide learning is variable: some students engaged in feedback to improve performance, whereas others did not.[4,9] Although clear communication of data via assessment dashboards is essential to provide feedback to various stakeholders, it should also mirror the CBME assessment schema.

### Methods of assessment visualisation

At the University of Michigan Medical School, we were particularly interested in how data visualisation could help create meaning for stakeholders. Specifically, students, faculty members and administrators were engaged in the design and prototyping. To support CBME assessment, we worked with our internal education software development team to build an assessment dashboard that could display and monitor each student's performance by pulling data from existing files. We used a Java platform. This required significant resources, including four software developers and two business analysts over 8 months, as well as biweekly to monthly meetings with a faculty member advisory team. A critical success factor in the development of the visualisations was the partnership between faculty members and information technology (IT) staff. When IT staff develop visualisations independent of faculty members, the prototypes may not be ideal for the intended purpose. When faculty members develop prototypes in isolation, the feasibility for IT implementation can be a limiting factor.

Historically, starting with first-year medical students, we provided a tabled summary of medical knowledge assessment scores. This tabular display included each student's individual quiz and examination scores as well as class performance metrics (i.e. class mean and standard deviation; Table 1), and was the standard method of displaying scores for both learners and faculty members.

Our tabular reporting of scores made it difficult to look at discrete metrics and understand how the student was performing longitudinally. Therefore, we asked

**Table 1. Standard method of displaying data**

**Scientific Trunk Course Cumulations** Download CSV

| Date | Course | Cumulative Score | Class Mean | Standard Deviation | Grade |
|---|---|---|---|---|---|
| 08/29/2016 | POP500 Genetics Component Cum 2016-2017 | 76.12 | 86.91 | 5.97 | S |
| 08/29/2016 | POP500 MDM Component Cum 2016-2017 | 87.05 | 87.51 | 8.4 | S |
| 08/29/2016 | POP500 Pathology Component Cum 2016-2017 | 76.52 | 91.38 | 6.76 | S |
| 10/03/2016 | CEL500 Course Cum 2016-2017 | 73.24 | 90.21 | 5.41 | S |
| 11/07/2016 | CAR504 Course Cum 2016-2017 | 85.24 | 89.21 | 5.47 | S |
| 11/14/2016 | REN506 Course Cum 2016-2017 | 83.53 | 87.40 | 5.04 | S |
| 12/19/2016 | MUS513 Course Cum 2016-2017 | 83.21 | 93.48 | 4.13 | S |
| 01/23/2017 | GAS508 Course Cum 2016-2017 | 83.70 | 88.59 | 4.60 | S |
| 02/13/2017 | END510 Course Cum 2016-2017 | 83.39 | 93.39 | 5.81 | S |
| 02/27/2017 | IMM501 Course Cum 2016-2017 | 89.55 | 86.21 | 5.27 | S |
| 03/27/2017 | CNS509 Course Cum 2016-2017 | 75.33 | 86.82 | 5.85 | S |
| 05/01/2017 | HPT512 Course Cum 2016-2017 | 81.23 | 94.88 | 5.23 | S |
| 05/08/2017 | INF500 Course Cum 2016-2017 | 83 | 85.55 | 4.87 | S |
| 05/29/2017 | CLN500/501 Course Cum 2016-2017 | 92.89 | 86.81 | 3.77 | S |
| 10/30/2017 | PSY614 Course Cumulation 2016-2017 | 83.22 | 87.06 | 4.21 | S |
| 11/25/2017 | NEU609 Course Cumulation 2016-2017 | 88.89 | 86.59 | 5.35 | S |
| 12/18/2017 | DER612 Course Cumulation 2016-2017 | 82.28 | 89.71 | 3.29 | S |
| 01/15/2017 | HEM603 Course Cumulative 2016-2017 | 94.17 | 94.43 | 6.23 | S |

Each student had access to a table as seen in the example with simulated student's scores and comparison scores of the class.

**... this graph – centred around the mean of the class – did not reflect our goal towards CBME assessment**

our education software development team to create a graphical representation of each student's performance across the first year of medical school (Figure 1a). This graph provided a clear picture of performance trends for each student, which was much more meaningful than the previous table of scores. In the example (Figure 1a), one can see that the student was below the class mean on the majority of the assessments. The first version of these graphs was norm-referenced, as it was anchored on the class mean and standard deviation.

As we looked at the graph in Figure 1(a), we appreciated how it told the story of the student's performance; but we also realised that this graph – centred around the mean of the class – did not reflect our goal towards CBME assessment.[1] In our competency committee, we wanted to shift from norm-referenced to more criterion-based assessment data for decision making. So, we went back to the software development team and addressed the question of how could we promote criterion- and competency-based decisions using data visualisation techniques?

The education software development team created a criterion-based visualisation for each student's performance (Figure 1b). The colour coding of each band derived its meaning from a traffic light: i.e. red, yellow and green. The criterion in the pre-clerkship basic science courses was established with a cut-off point of 75%. A score below 75% was in the red zone; this signalled that the student had failed a course and was not competent in the domain. Although 75% was considered a pass, this score signalled that a student was barely performing at criterion levels. We knew from historical data and analyses that students with 2-year pre-clinical cumulative scores below 83% were at increased risk of failing the United States Medical Licensing Examination (USMLE) Step 1 examination, so the band from 75% to 83% was coded yellow to indicate this risk and marginal competence. These indicators helped us to identify, communicate and work with students to improve their performance. Scores ranging from 83% to 91% were coded green, indicating that students were solidly competent. We were concerned that a competency-based system might focus only on the lowest performers (the floor); therefore, we also noted high performance or 'excellence' based on students achieving scores above 91%.

We deliberated about whether to include any norm-referenced data in our criterion graph.[1] If we believed in a purely competency-based system, it would not matter how the other students performed and there would be no need to display the mean; however, we ultimately decided to include the class mean on the graph (plotted with triangles) because we felt

**... it is important to pay attention to the alignment with the construct of competency and to engage stakeholders ...**

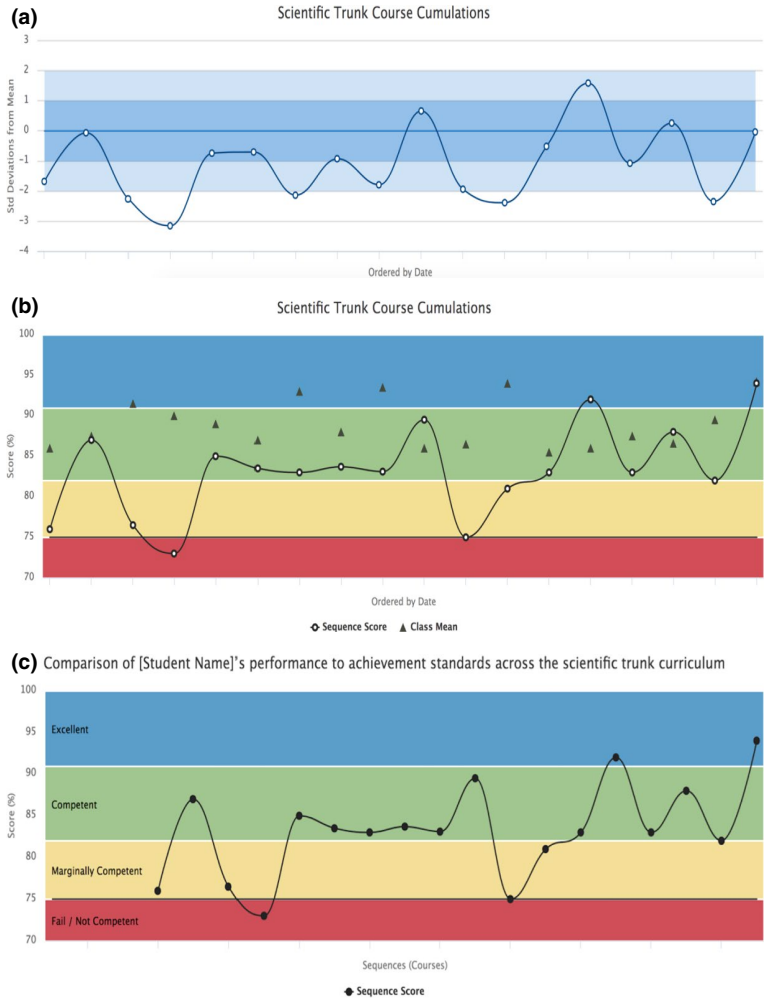



**Figure 1.** Data visualisations. (a) Norm method of displaying data: The dashboard has each student's performance centred on the mean of the class for each exam at 0 with standard deviation around the mean. Performance of this simulated student was low in the beginning and improved over time. (b) Criterion method of displaying data: The competency performance bands are the visually dominant element on the page, centring the students' performance. The line connecting each point of performance data assisted in communicating trends (up or down in performance) as well as giving the simulated student's performance scores more visual weight in the visualisation. The class mean is noted as well. (c) Criterion method of displaying data without norm-references. The competency performance bands are the visually dominant element on the page, centring the students' performance. The line connecting each point of performance data assisted in communicating trends (up or down in performance) as well as giving the simulated student's performance scores more visual weight in the visualisation.

that including norm-referenced data provided meaningful knowledge to the student about performance (Figure 1c). This decision was grounded in historical data: the yellow zone was associated with an increased risk for failing Step 1. Kruger and Dunning have shown that in general people believe that they are above average, even when their performance is below average.[10,11] This self-assessment flaw is also seen in students, so providing class statistics addresses this limitation.

Our data visualisation has multiple purposes, including providing formative feedback to students, providing a platform for coaching and counselling students, and informing decision making. Students review their performance data with coaches or advisors to identify areas of strength and further growth and goal setting.[9,11,12]

### Stakeholder feedback

We found it important to have faculty member and assessment expertise in the development process. Our team comprised of software developers as well as experts in educational assessment and theory and faculty members from core content areas. Together, we explored multiple iterative visualisations. Once the final version was created, it was shared with the educational leadership faculty members and the competency committee who appreciated the clarity. In contrast, the students questioned the decision to include four categories instead of just three (i.e. red, yellow and green only). Students said they were 'trying to figure out how the cut-offs were determined since students will likely want to know'. Although the marginal (yellow) cut-off score was based on advisory historical data, it was not predictive of poor performance. We had less evidence to support the 'excellent' category cut-off score, but we felt that it was important to note excellence. In the process of implementation, we recognised the importance of communicating goals to our students.

In our clerkship assessment dashboard, we used a purely competency-based display of data without norm-referenced metrics (Figure 2). As our journey towards CBME assessment progresses, we continue to explore the challenges surrounding the standard setting for cut-off scores to indicate competency in medical knowledge. Furthermore, we hope to be able to remove all norm-based references. Yet this will require a culture shift towards a growth mindset and away from competition and the rank ordering of students.

## RECOMMENDATIONS FOR CBME ASSESSMENT DATA VISUALISATION

During this journey, the CBME assessment framework helped to guide the development and visualisation of student scores and to create meaning. Boscardin and colleagues' 12 tips to promote successful development of a learner performance dashboard provides excellent guidance.[7] There were lessons learned during

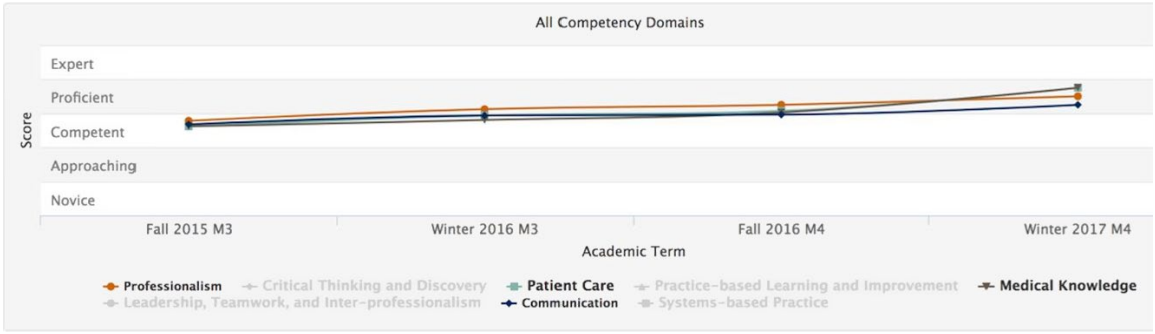



**Figure 2.** Competency-based clerkship dashboard for four competencies over the year. Note: The competency performance of four competencies of a simulated student is graphed over time of the third- and fourth-year of medical school. Competency bands (competent, proficient, expert) are demonstrated visually. The line connecting each point of performance data assisted in communicating trends (up).

## Table 2. Tips to data visualisation[7]

| Tip | Application |
|---|---|
| What was your assessment framework? | Base dashboard development on an assessment framework and ensure that your data visualisation reflects it |
| Who did you engage in the development? | Form a dashboard team with education leaders, information technology experts, assessment experts, data managers and learners. Bring in all stakeholders including students in the design and implementation process |
| Did you identify and display performance benchmarks? | Determine which benchmarks have meaning and indicate them. Focus on the clarity of representing the data |
| Did you emphasise your points in the chart? | Choose a chart type that best fits your data and add features that emphasise your point (e.g. colors, comparison data, highlights and/or annotations of particular data points) and consolidate key performance metrics using at a glance data visualisations |
| Did you focus on the purpose of the data visualisation? | Create a descriptive headline that explains to users what they will see when they look at the chart |

the development of this assessment dashboard (Table 2).

## CONCLUSIONS

When designing a reporting data visualisation to abstract, summarise or categorise assessment data for a CBME assessment programme, it is important to pay attention to the alignment with the construct of competency and to engage stakeholders in the process.

### REFERENCES


1. Pereira AG, Woods M, Olson APJ, van den Hoogenhof S, Duffy BL, Englander R. Criterion-based assessment in a norm-based world: how can we move past grades? *Acad Med* 2018;**93**(4):513–515.
2. Hauer KE, Lucey CR. Core clerkship grading: the illusion of objectivity. *Acad Med* 2019;**94**(4):469–472.
3. Van der Vleuten CP, Schuwirth LW, Driessen EW, Dijkstra J, Tigelaar D, Baartman LK, van Tartwijk J, A model for programmatic assessment fit for purpose. *Med Teach* 2012;**34**(3):205–214.
4. Monrad SU, Mangrulkar RS, Woolliscroft JO, Daniel MM, Hartley SE, Highet A, Vijayakumar N, Santen SA. Competency committees in undergraduate medical education: approaching tensions through a polarity framework. *Acad Med* 2019. https://doi.org/10.1097/ACM.0000000000002816. [Epub ahead of print]
5. Friedman KA, Raimo J, Spielmann K, Chaudhry S. Resident dashboards: helping your clinical competency committee visualize trainees' key performance indicators. *Med Educ Online* 2016;**21**:29838.
6. Spickard A III, Ahmed T, Lomis K, Johnson K, Miller B. Changing medical school IT to support medical education transformation. *Teach Learn Med* 2016;**28**(1):80–87.
7. Boscardin C, Fergus KB, Hellevig B, Hauer KE. Twelve tips to promote successful development of a learner performance dashboard within a medical education program. *Med Teach* 2018;**40**(8):855–861.
8. Rost LC. *What Questions to Ask When Creating Charts: The Attempt of a Data Vis Workflow*. *Chartable*. Available at https://blog.datawrapper.de/better-charts/. Accessed on 5 November 2017.
9. Hauer KE, Iverson N, Quach A, Yuan P, Kaner S, Boscardin C. Fostering medical students' lifelong learning skills with a dashboard, coaching and learning planning. *Perspect Med Educ* 2018;**7**(5):311–317.
10. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol* 1999;**77**(6):1121–1134.
11. Sargeant J, Eva KW, Armson H, Chesluk B, Dornan T, Holmboe E, Lockyer JM, Loney E, Mann KV, van der Vleuten CP. Features of assessment learners use to make informed self-assessments of clinical performance. *Med Educ* 2011;**45**(6):636–647.
12. Vance G, Williamson A, Frearson R, O'Connor N, Davison J, Steele C, Burford B. Evaluation of an established learning portfolio. *Clin Teach* 2013;**10**(1):21–26.

**Corresponding author's contact details:** Sally A Santen, Department of Emergency Medicine, VCU School of Medicine, Virginia Commonwealth University, Richmond, Virginia 23284-2512, USA. E-mail: Sally.santen@vcuhealth.org

**Funding:** None.

**Conflict of interest:** The University of Michigan Medical School and Virginia Commonwealth University obtain funding for the Accelerating Change in Medical Education Grant from the American Medical Association.

**Acknowledgements:** None.

**Ethical approval:** This was not a research study; therefore, ethical approval was not obtained.

doi: 10.1111/tct.13106