The Harvard Medical School Pathways curriculum: A comprehensive curricular evaluation

Amy M. Sullivan, Edward Krupat, Jules L. Dienstag, Jakob I. McSparron, Stephen R. Pelletier, Sara B. Fazio, Thomas J. Fleenor, John L. Dalrymple, Edward M. Hundert, Richard M. Schwartzstein

Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

Purpose: The Harvard Medical School Pathways curriculum represents a major reform effort. Our goals were to enhance reasoning and clinical skills and improve the learning environment and students’ approach to learning via use of collaborative, case-based pedagogy; early clinical exposure; and enhanced approaches to teaching and evaluating clinical skills. We evaluated the impact of Pathways on key outcomes related to these goals.

Materials and methods: In this prospective, mixed-methods study, we compared the last prior-curriculum cohort (2014 matriculation, n = 135) and first new-curriculum cohort (2015 matriculation, n = 135). Measures included Likert-type surveys, focus groups, and test scores to assess outcomes.

Results: Compared with prior-curriculum students, new-curriculum students reported higher mean preclerkship learning environment ratings (Educational Climate Inventory, 62.4 versus 51.9, p < 0.0001) and greater satisfaction with the quality of their preclerkship education (88% versus 73%, p = 0.0007). Mean USMLE Step-1 and Step-2 scores did not differ between groups. At graduation, new-curriculum students rated their medical school experience higher in 6 of 7 domains, including ‘fostering a culture of curiosity and inquiry’ (4.3 versus 3.9, p = 0.006) and focus on ‘student-centered learning’ (3.9 versus 3.4, p = 0.002).

Conclusions: The new curriculum outperformed or was equal to the prior one on most measures of learning environment and perceived quality of education, without a decline in medical knowledge or clinical skills. Robust longitudinal evaluation provided important feedback for ongoing curriculum improvement.

Introduction

Despite calls for more rigorous curriculum-related research (Gruppen 2007; Fincher et al. 2010; Schwartzstein 2017), evidence for effectiveness of newly unveiled curricula is often limited in scope. Although all schools strive to assess the effectiveness of major curricular reforms, such whole-curriculum outcome studies involve methodologic and logistic challenges, and also require significant resources at a time when available resources are usually stretched to deal with the curriculum change itself. Consequently, other than efforts to compare United States Medical Licensing Examination (USMLE) scores and/or student satisfaction levels between old and new curricula (Heiman et al. 2018), few comprehensive program evaluations have been reported (Karpa and Abendroth 2012; Fischel et al. 2019).

In 2015, Harvard Medical School (HMS) Harvard Medical School Medical Education (2021) introduced the Pathways curriculum (Schwartzstein et al. 2020), a comprehensive curriculum reform to enhance the learning environment, student curiosity and readiness to learn with an emphasis

Practice points

- A Case-Based Collaborative Learning approach, which promotes active learning, curiosity, growth mindset, student- and patient-centered learning, and attention to the quality of the learning environment can improve student satisfaction with their medical school experience while maintaining academic rigor.
- A shortened preclerkship period and earlier clinical exposure may require ongoing monitoring to assess students’ needs for additional clinical preparation.
- Commitment to evidence-based medical education efforts requires planning and dedicated resources to evaluate the effectiveness of a new curriculum.
- Multiple methods to assess the impact of a new curriculum, including student and faculty perspectives, can provide important early indicators of the curriculum’s impact on learning.
- Ongoing quality improvement efforts are critical for identifying and allowing for rapid responses to address weaknesses in a new curriculum.

CONTACT A.M. Sullivan asulliv5@bidmc.harvard.edu Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

Supplemental data for this article can be accessed online at https://doi.org/10.1080/0142159X.2022.2081142

© 2022 Informa UK Limited, trading as Taylor & Francis Group
on active learning, critical thinking, and growth mindset (Ricotta et al. 2019); a new model of collaborative, case-based pedagogy (Krupat et al. 2016); earlier clinical exposure; an enhanced advising system; and revised teaching and evaluation of clinical skills. HMS is an urban undergraduate medical school located in Boston, Massachusetts, US. There are 12,209 faculty who teach, conduct research, and provide clinical care at one of fifteen hospital affiliates. There are approximately 750 medical students in the school, about half of whom spend one or more extra years either doing research or getting a second degree (master’s or PhD). Fifty-one percent of the students are women, and 23% are students from racial and ethnic groups that are historically underrepresented in medicine (URiM).

The details of the development of this curriculum are provided elsewhere (Schwartzstein et al. 2020). Briefly, the call for a new curriculum came from the faculty, who desired to change the pedagogical approach, particularly within the preclerkship courses, to increase student interactions, peer teaching, and emphasis on analytical reasoning based upon a strong foundation of biomedical concepts. Thinking is emphasized over factual memorization; diseases are examined in close proximity to teaching of underlying anatomy and physiological principles with diagnoses as exemplars of human biology rather than as illness scripts. Content transfer is accomplished largely with a flipped classroom approach; lectures are minimized. Preclerkship case-based collaborative learning (Krupat et al. 2016) is the mainstay of small group learning in the new curriculum, occupying the vast majority of student-faculty contact time. Clinical experiences are introduced in the first weeks of matriculation, including a longitudinal primary care experience; the core clerkship year begins after 14 months of instruction with clinical skills assessment anchored to entrustable professional activities (EPAs). Advanced integrated basic science courses, addressing cutting edge discoveries linked to patient care, are required in the post-clerkship phase of the curriculum.

In parallel with the development of the Pathways curriculum, we initiated a comprehensive program evaluation. We addressed the following research questions: Compared with students from the prior curriculum, do students in the new curriculum demonstrate (a) better levels of self-reported attitudes such as curiosity, patient-centeredness, and perceptions of the learning environment; (b) better perceptions of the overall quality of their medical education; and (c) equal or better levels of knowledge and/or clinical skills? In addition to assessing the new curriculum, we committed to an ongoing quality improvement approach to curriculum development, consistent with the philosophy that evaluation should be a process rather than an event (Konopasek et al. 2016). Our study aimed to offer evidence for effectiveness and a compendium of ‘lessons learned’ from our efforts to conduct a comprehensive curriculum evaluation.

**Materials and methods**

Because a randomized controlled trial was not feasible, we compared the first entering cohort of the new curriculum to historical controls, the last entering cohort of the prior curriculum, using both quantitative and qualitative methods. We also conducted within-group, longitudinal quality improvement assessments (Wong and Headrick 2021) that focused on the new curriculum students as they progressed through clerkship to graduation. This prospective mixed methods approach (Creswell and Plano Clark 2017) included input from students and faculty, provided a rich and holistic characterization of the strengths and weaknesses of the curriculum, and provided feedback allowing rapid response to any identified weaknesses. Overall, we used a hybrid evaluation approach (Frye and Hemmer 2012) consistent with the Kirkpatrick New World Model (Kirkpatrick and Kirkpatrick 2021) and the principles of the Joint Committee for Standards on Educational Evaluation (Yarbrough 2010). The HMS Institutional Review Board characterized the study as educational quality improvement, not subject to review as human-subjects research.

**Participants**

Participants were HMS students who entered the prior curriculum in August 2014 (n = 135) and the new curriculum in August 2015 (n = 135); Step 1 scores from the 2016 entering class were also included in the analysis. Students in the Harvard-Massachusetts Institute of Technology Health Sciences and Technology program (HST), who take a different preclerkship science curriculum, were excluded. Clinical clerkship directors also completed a survey concerning the clinical skills of the two cohorts. For the focus groups, we sent emails to all students in the prior curriculum and the first cohort of the new curriculum (excluding HST students) inviting them to participate.

**Outcomes and measures**

As summarized in Table 1, we collected a broad range of outcome measures.

**Attitudes and learning environment**

We hypothesized that, compared with students in the prior curriculum, new-curriculum students would rate the learning environment more positively, be more satisfied, and have greater levels of curiosity and empathy, with similar or reduced burnout at the end of their preclerkship curriculum.

Matching students’ baseline scores within a month of matriculation to their end-of-preclerkship responses via electronic survey, students completed several instruments with strong validity evidence: the Jefferson Scale of Empathy (JSE) (Tavakol et al. 2011); Patient-Practitioner Orientation Scale (PPOS) (Krupat et al. 2000), a measure of patient-centered beliefs; Need for Cognition Scale (NFCS) (Cacioppo et al. 2013), which assesses student predisposition toward effortful thinking; and Litman’s Curiosity Scale (LCS) (Litman et al. 2010). At the end of the preclerkship curriculum, we also assessed student experiences of the learning environment through the Educational Climate Inventory (ECI) (Krupat et al. 2017), which measures the extent to which the culture for learning is perceived to be performance-oriented versus growth/mastery-oriented.
We drew on existing data from the Association of American Medical Colleges (AAMC 2015–2016) Medical School Year Two Questionnaire (Y2Q)(2015–2016), selecting items covering satisfaction with medical school experience; quality of the learning environment, as measured by the Medical School Learning Environment Survey (MSLES) (Smith et al. 2016); and student well-being, as measured by the Oldenburg Burnout Inventory (OLBI) (Halbesleben and Demerouti 2005).

To assess students’ overall perceptions of their satisfaction with their medical education at the completion of their training, we developed the HMS Exit Survey, administered to graduating cohorts of the prior and new curricula. We chose survey domains based on a review of the literature and HMS-specific objectives and curricular elements. Specifically, the survey addressed student satisfaction with the curriculum, advising, student assessment, and core clerkship and post-core clerkship experiences. We based survey domains on learning and learning environment objectives and conducted cognitive interviews with students not involved in the study to refine the instrument (Artino et al. 2014).

We complemented our quantitative data collection with information obtained from focus groups, conducted between March and October 2016, when both cohorts were in the early phases of their clerkships. We developed the focus group guide based on a review of literature and specific aims of the new curriculum. Focus groups explored student responses to questions about the curriculum’s support for the development of critical thinking, relationships with peers and course faculty, most and least valuable sources of learning, and the culture for learning. Focus groups with students from both cohorts were audio-recorded, transcribed, and de-identified. One of the authors (AMS), a social scientist who had no supervisory relationships with students, led the focus groups, which were conducted in private conference rooms in the medical school. A copy of the focus group guide is shown on Supplementary Appendix A.

### Knowledge and performance

Using results obtained from the National Board of Medical Examiners (NBME) USMLE Step 1 and Step 2 Clinical Knowledge (CK), we hypothesized that students in the new curriculum would demonstrate equal or better performance in these areas as a result of the greater emphasis on teaching methods that supported learning and long-term knowledge retention (Schwartzstein et al. 2020).

---

**Table 1. Summary of measures used to compare student cohorts and faculty perceptions of student knowledge, skills, and attitudes in the prior and new curricula.**

| Category and New World Kirkpatrick (NWK) Level | Constructs | Instruments | Timing of administration |
|-----------------------------------------------|-------------|-------------|--------------------------|
| Attitudes and perceptions of learning environment | Empathy | Jefferson Empathy Scale (JSE), a 20-item self-report instrument (pre- and post). | A,B |
| | Patient-centeredness | Patient-Practitioner Orientation Scale (PPOS), an 18-item survey instrument measuring patient-centered beliefs. | |
| NWK 1, 2, 4 | Need for cognition | Need for Cognition Scale (NFCs), 18 items measuring predisposition toward effortful thinking | |
| | Curiosity, defined as an intrinsically-motivated drive for learning and information-seeking | I-Type and D-Type Scales, 5 items each | |
| Learning environment | Educational Climate Inventory, 20-item instrument with three subscales: Centrality of learning and mutual respect; competitiveness and stress; and Passive learning and rote memorization. | B |
| | AAMC Year Two Questionnaire, annual survey for 2nd year students, includes learning climate items from the Medical School Learning Environment Survey (MSLES) and 8-item Oldenburg Burnout Inventory | D |
| Student satisfaction with medical education | Focus groups | Exit Surveys | B |
| Knowledge NWK 2 | Foundational medical knowledge | USMLE Step 1 Exams. Medical content knowledge in anatomy, behavioral sciences, biochemistry, biostatistics and epidemiology, microbiology, pathology, pharmacology, and physiology. | B |
| Skills NWK 3 | Core clinical skills-summative | 9-Station summative Objective Structured Clinical Examination (OSCE) (at beginning of final year). Summative assessment of observed student skills in conducting history, physical exam skills, communication and clinical reasoning in standardized test setting | E |
| | Student clinical readiness | Clerkship director faculty perceptions of each student cohort’s preparation to learn and participate in the clinical environment. 11-item survey with ratings for students from each of the two cohorts. Items covered observed knowledge, physical examination and presentation skills, and openness to feedback. | C |
| | Clinical skills—formative, longitudinal | Entrustable professional activities focused on 13 specific skills. Longitudinal, formative assessment of student clinical skills across different phases of the curriculum, including pre-clerkship, clerkship and post-clerkship courses. | G |

*Key: A = Administered at start of preclerkship period (Year 1); B = End of preclerkship period; C = First quarter of first year of clerkship; D = Mid-year 2nd year; E = Final year before graduation; F = Graduation; G = Administered across entire curriculum* 

*New World Kirkpatrick (NWK) model levels: 1. Reaction (satisfaction, engagement), 2. Learning (attitudes, knowledge, skills), 3. Behaviors (application of learning), and 4. Leading indicators (here, perceived learning environment, curiosity, empathy, and burnout).*

*USMLE: U.S. Medical Licensing Exam (3 years of data were used for Step 1, 2 years of data available for Step 2 CK)*
Clinical skills and observed behaviors

To assess faculty perceptions of students’ readiness for clerkships in the context of the shortened preclerkship phase, we surveyed clerkship directors at all HMS clinical sites three months after the arrival of new curriculum students to determine perception by clerkship directors of student clerkship readiness in skills, knowledge, and clinical reasoning abilities using a 5-point scale (1 = not very ready, 5 = extremely ready). Clerkship directors were asked to rate their general impressions of student readiness for each cohort as follows: ‘As clerkship directors you have had the opportunity to work with and receive feedback about many cohorts of students in the New Pathway (the previous HMS curriculum) and now with a first cohort of students in the Pathways Curriculum (the new curriculum). Please rate your general impressions of students in these two groups as to their readiness to contribute in their very first rotations of the core clerkship year.’

Once students entered their clerkship curriculum, we monitored clinical progress of students in by tracking entrustable professional activities (EPAs), creating a longitudinal OSCE (objective structured clinical examination) program, and instituting the Clinical Capstone course (a required, culminating clinical experience for fourth-year students), each described elsewhere (Schwartzstein et al. 2020). The 13 EPAs, modified from the AAMC framework (Obeso et al. 2017), were mapped across all clinical courses in the curriculum (Supplementary Appendix B). We monitored longitudinal student progress across the clerkship year and other phases of the curriculum. Because the EPAs were introduced as part of the new curriculum, we did not have comparative data from students in the prior curriculum.

Analytic approach

Statistical analyses were carried out with JMP® Pro Version 14 (SAS Institute Inc., Cary, NC, USA, 1989–2019) and SPSS Version 25 (IBM SPSS Statistics for Windows, Armonk, NY, USA, 2017).

For comparison of pre- and post-survey attitudes between the two cohorts, we carried out mixed models analyses (Detry and Ma 2016) testing for interactions between curriculum (prior and new) and time (pre and post). For comparison of multiple means (e.g., Step 1 exams over three years), we conducted analyses of variance. We tested differences between groups on categorical items with chi-square tests of independence and tested mean differences with independent groups t-tests. We set our alpha level at 0.05 (two-tailed) and addressed the issue of multiple comparisons by focusing our primary outcomes on a few overarching domains, each representing a single underlying construct (attitudes, learning environment, knowledge, and clinical skills) (Schochet 2008). Effect sizes are reported as Cohen’s d for t-tests and eta-squared for analyses of variance.

For student focus group data, we conducted qualitative analysis by applying the Framework method of content analysis (Ritchie et al. 2013). One investigator (AMS) and a research coordinator analyzed focus group transcripts weekly, beginning with discussions of close readings of each transcript, formulation of a codebook, and coding of data with software Dedoose® (Version 8, www.dedoose.com, SocioCultural Research Consultants, LLC, Los Angeles, CA). Differences in coding and interpretation were resolved by consensus. Trustworthiness of results was supported through presentations and discussions of results with the larger research team, identification of saturation of themes, and comparison with quantitative findings in surveys (Shenton 2004). We used the Standards for Reporting Qualitative Research (SRQR) (O’Brien et al. 2014) to guide our design and reporting of findings.

Results

Attitudes and learning environment

We found no significant baseline nor pre-post differences on our measures of empathy, patient-centeredness, curiosity, or need for cognition. Because response rates for the post-surveys were lower than those for pre-surveys (ranging from 16.3–52.6%), we carried out follow-up tests for differences between students who had paired (pre-post) data and those who had only pre-survey data. We found no statistically significant differences between groups on any of these measures. At the end of the preclerkship period, new-curriculum students reported higher average ratings on the Educational Climate Inventory compared with the prior-curriculum students’ scores (62.4 versus 51.9, respectively, p < 0.0001, ES = 1.1), suggesting a greater orientation in the new curriculum toward learning/mastery.

The Y2Q was distributed during the first quarter of the core clerkship year for new curriculum students, and showed more positive attitudes among students in the new curriculum compared with those in the prior curriculum, with most differences representing moderate or large effect sizes (Table 2). New-curriculum students reported higher levels of satisfaction with the quality of their medical education and rated the learning environment more positively. On the Oldenburg Burnout Inventory, they reported lower levels of burnout along the ‘disengagement’ dimension. No differences were observed between groups on the ‘exhaustion’ dimension.

A total of 15 students participated in four focus groups, ranging from 2-5 students in each group: 10 from the new curriculum and five from the prior curriculum (see detailed results in Supplementary Appendix C). Students in the new curriculum valued highly the ‘flipped classroom’ approach, liked the structure that helped them come to class prepared, and were enthusiastic about opportunities to engage with other students in discussing cases and problems. The few complaints voiced by these students reflected the shortcomings of some sessions perceived as not fully developed in the first year of the new curriculum. Despite four fewer months of course time prior to the start of their clerkships, new-curriculum students felt as prepared to work in the clinical environment, particularly in their ability to communicate with patients, conduct a history and physical exam, present patients, and think critically about disease processes and mechanisms compared with prior-curriculum students. Students in the new curriculum, however, felt less prepared in the treatment and management of patients and in having a strong knowledge.
base in anatomy, drug treatments, and correct procedures for ordering and interpreting clinical tests.

At graduation, students in the new curriculum (Figure 1) reported higher average ratings on six of seven Exit Survey items related to the overall medical school experience, all with medium effect sizes. New-curriculum students rated their curriculum more highly in fostering a culture of curiosity and inquiry; relying upon innovative and effective pedagogical models; promoting a partnership between faculty and students; focusing on a patient-centered approach to care as well as a student-centered approach to learning; and promotion of habits of life-long learning. No statistically significant differences were present in students’ perception of their training in critical thinking skills.

Knowledge and performance

On USMLE Step 1 exams, mean scores of the students in the prior and new curriculum were not significantly different (prior curriculum 247.4, first and second cohorts of new curriculum 244.6 and 248.6, respectively, p = 0.06). Sub-scale scores of the Step 1 exam, however, indicated that, for students in the new curriculum, the tail of the distribution was shortened at the low end; i.e. students scoring below the class mean did better with the new curriculum (Figure 2). Step 2 CK exam results showed no statistically significant differences between means for the prior (252.4) and new (253.5, p = 0.56) curriculum groups.

In surveys to assess students’ clinical preparedness in the early clerkship phase, clerkship directors rated new-curriculum

Table 2. AAMC Year 2 Questionnaire (Y2Q) results for HMS students in the prior and new curriculum groups: satisfaction with education quality, learning environment, and burnout, with comparative data for all US medical schools.

| Scale/Item                            | All medical schools | Prior curriculum (2015) | New curriculum (2016) | p-Value, effect size |
|---------------------------------------|---------------------|-------------------------|------------------------|----------------------|
| Satisfaction with quality of medical education (%) | 85.4% | 73.0% | 88.1% | p = 0.007 |
| ‘agree’ or ‘strongly agree’ | 12,405 | 126 | n = 101 |
| Learning Environment Scales* | | | | |
| Emotional climate | 9.2 (3.1) | 8.3 (3.0) | 9.8 (2.7) | p = 0.0003, ES = 0.50 |
| n = 11,627 | n = 116 | n = 89 |
| Student-faculty interactions | 14.7 (3.2) | 13.1 (3.1) | 15.5 (2.7) | p < 0.0001, ES = 0.77 |
| n = 11,646 | n = 118 | n = 92 |
| Student-student interactions | 14.9 (3.2) | 14.3 (3.2) | 16.0 (2.9) | p < 0.0001, ES = 1.25 |
| n = 11,626 | n = 118 | n = 91 |
| Oldenburg Burnout Scale | | | | |
| Disengagement | 9.7 (3.7) | 10.2 (4.2) | 9.0 (3.4) | p = 0.0287, ES = 0.29 |
| n = 11,145 | n = 120 | n = 88 |
| Exhaustion | 11.7 (3.8) | 11.6 (4.1) | 10.7 (3.7) | p = 0.11 |
| n = 11,066 | n = 114 | n = 85 |

p-Values represent tests for differences between students in prior and new curriculum groups.

Overall, my medical education at HMS...

Fostered a culture of curiosity and inquiry. (P = 0.0057, ES = 0.67)

Utilized innovative and effective pedagogical models. (P < 0.0001, ES = 0.67)

Promoted a partnership between faculty and students. (P = 0.0007, ES = 0.57)

Focused on a student-centered approach to learning. (P = 0.0018, ES = 0.53)

Focused on a patient-centered approach to care. (P = 0.0032, ES = 0.48)

Promoted habits of life-long learning. (P = 0.0218, ES = 0.38)

Enhanced my critical thinking skills. (P = 0.06)

Figure 1. Student responses to 2018 and 2019 graduating year Harvard Medical School Exit Surveys, from students who matriculated in 2014 or earlier (prior curriculum, n = 128) and students who matriculated in 2015 (current curriculum, n = 53). (Results shown as mean scores).
students more positively than prior-curriculum students in openness to feedback (4.8 versus 4.4, \( p = 0.003, ES = 0.20 \)) and inquisitiveness/curiosity (4.7 versus 4.4, \( p = 0.001, ES = 0.16 \)). However, they rated new-curriculum students lower in breadth (3.3 versus 4.2, \( p < 0.001, ES = 0.42 \)) and depth of knowledge (3.6 versus 4.2, \( p = 0.004, ES = 0.25 \)), presentation (3.6 versus 4.1, \( p = 0.005, ES = 0.29 \)), and physical exam skills (3.4 versus 3.8, \( p = 0.005, ES = 0.19 \)). No differences were detected in ratings of problem-solving, interviewing skills, persistence in finding answers, teamwork skills, or motivation to learn.

In the first cohort of new-curriculum students who graduated in 2019, all students completed the Clinical Capstone; faculty documented achievement of all EPAs at or close to the entrustable level (level 4 out of 5 required for graduation). The graduating class of the new curriculum achieved a 100% first-attempt pass rate on Step 2 Clinical Skills (CS,) compared to 95% in the last class of the prior curriculum (both classes were subjected to the increased passing standards set for the Step 2 CS exam in 2017).

**Discussion**

In this prospective analysis of the first four years of the HMS Pathways curriculum, we addressed a broad set of outcomes related to the goals of the curriculum reform – students’ engagement and perceptions of the learning environment, preparation for classwork, growth mindset, emphasis on thinking skills, depth of understanding, and overall quality of their medical school education. On the vast majority of the measures of the student experience and approach to learning, the new curriculum outperformed its predecessor, an outcome achieved without a decline in medical knowledge or academic performance as indicated by USMLE Step 1 and Step 2 CS.

Students in this first Pathways cohort performed as well as the last cohort of the previous curriculum as they
progressed through their clerkship and postclerkship years. In fact, subscore Step 1 measures suggested improvement for the lower half of the class, which was consistent with our curriculum pilot studies (Krupat et al., 2016). Early assessment of clinical skills showed some deficiencies based on clerkship director surveys, but the lower assessment of student ability to elaborate a complete differential diagnosis was not unexpected since the new preclerkship curriculum focuses on exemplar diseases to illustrate pathophysiological principles; clinical reasoning is approached from the standpoint of pathophysiology rather than illness scripts; and analytical reasoning is emphasized rather than pattern recognition, which likely slowed the process of creating differential diagnoses among novice learners. Nevertheless, we believe these skills will reduce cognitive biases in future years and reduce diagnostic error (Royce et al. 2019). The effect sizes of these differences were small and the identified gaps were addressed quickly through curricular enhancements made for subsequent cohorts.

Our goals for improved student experiences of learning and the learning environment were validated, with all differences representing large effect sizes. Perceptions of the new preclerkship phase of the curriculum reflected greater support of mastery rather than performance orientation as measured by the ECI. The Y2Q lower burnout subscores in disengagement and the absence of a difference in exhaustion subscales indicated the new compressed preclerkship curriculum was not posing an excessive burden on students. Both the improved learning environment and the enhanced advising system may have contributed to the lower burnout scores.

Reports from the core clerkship directors indicated that new-curriculum students were more curious and open to feedback compared to the prior-curriculum cohort. In our Exit Survey, students from the new curriculum provided more positive ratings of qualities we sought to nurture, such as curiosity and adhering to both patient-centered and student-centered learning.

Clinical skills (specifically, physical exam and interviewing) were not as strong at the end of the shortened preclerkship phase in the new compared to the prior curriculum. Given the dedication of one full weekday to clinical skills throughout the preclerkship phase, we were surprised by this finding. In response, we made changes in the teaching of these skills for the second cohort of students in the new curriculum, refined our use of OSCEs and other tools for assessment of clinical skills, and developed a coaching program for students who underperformed clinically. Our experience with the required Clinical Capstone and our 100% first attempt passing rate on Step 2 Clinical Skills (CS) provide independent evidence that students in the first Pathways group achieved the requisite clinical skills prior to graduation.

Limitations

In studying the first year of the new curriculum, we encountered several challenges, which constitute study limitations; these may provide guidance for those contemplating evaluations of curricular change. The first year of any new curriculum, particularly one in which all the preclerkship courses are reconstructed entirely and novel pedagogy adopted, is never perfect. Some curricular content, although elaborately planned and vetted, was not finalized until shortly before deployment. Because faculty implemented many changes for the second year of Pathways, an analysis of the first cohort of Pathways students may have underestimated its ultimate value and the subsequent learning outcomes.

Great excitement accompanied the unveiling of the new curriculum, which had the unintended consequence of alienating some students in the last class of the prior curriculum and may have contributed to lower response rates than desired from that class on the assessments we developed to evaluate the new curriculum. Our concern, however, was offset by the alignment of the survey findings with focus group data as well as by the lack of differences (in pre-survey responses) between post-survey respondents and nonrespondents. The prior-curriculum cohort also had a low ECI response rate; we were able, however, to conduct post hoc tests comparing the mean for this group with that of HMS second-year students in a prior ECI validation study (with a 51.5% response rate) (Krupat et al. 2017). Again, we found no statistically significant results between groups, suggesting that our results still provided a fair representation of the prior-curriculum class. Finally, analyses of those students for whom we had both pre and post data on our other attitudinal measures (empathy, patient centeredness, and curiosity) indicated that the students were generally equivalent across cohorts upon entry into the curriculum and thereafter. Because students and faculty could not be blinded to the intervention, perceived differences in quality of the prior and new curriculum could be affected by bias among both groups, which remains a limitation of the study design (e.g., the possibility that hospital-based faculty less involved in the reform may have had a bias that a shortened preclerkship curriculum could not possibly have prepared students as well). This threat to the validity of our results is offset, to the extent possible, by our triangulation of data sources and the longitudinal tracking of the progress of students in the new curriculum.

Methodologically, efforts to assess a new curriculum that was still in the process of being refined provided a number of other challenges, including (1) resistance on the part of students to participate in program evaluations; (2) difficulty in adapting and creating instruments to assess critical thinking skills; and (3) difficulty in comparing clinical skills while simultaneously attempting to improve OSCEs and evaluation rubrics and establishing a cadre of assessment faculty. We did have several advantages that supported our analysis of the new curriculum. Funding for the first phase of the evaluation (preclerkship and early clerkship) from the Harvard Institute for Learning and Teaching was invaluable in providing us with resources to design and implement in-house measures of attitudes, knowledge, and clinical skills and to collect data to compare the two cohorts. In addition, availability of school-level AAMC data such as the Y2Q survey and NBME Step 1 and Step 2 CK data—particularly with the availability of specific scores and subscores—were valuable in providing objective measures with national comparisons.

Conclusion

Among the major goals of our new curriculum were improvement in the learning environment and students’ perception of their readiness to learn, question, and grow as physicians. The
quantitative and qualitative data presented demonstrate that these goals were largely achieved without compromising student knowledge and clinical skills. We hope that other schools embarking on major curricular reforms will be motivated by and will learn from our experience to engage in efforts to rigorously evaluate pedagogical and organizational initiatives.

As knowledge of cognitive and learning sciences evolves, and as our healthcare system continues to change, medical school curricula must adapt as well. As medical educators, we must pledge to test rigorously our new approaches to and methods for teaching the next generation of physicians; we have a responsibility to ensure that we are providing the best evidence-based education and to report all outcomes, including those that do not initially go as well as expected. We have reported a comprehensive analysis of a new curriculum with successes and shortcomings; the Pathways curriculum is fulfilling most of our expectations, particularly with respect to our focus on thinking skills and providing an environment that engages students and is supportive of a learning orientation. Still, ‘a good curriculum is never finished.’ Consistent with our commitment to educational continuous quality improvement, we continue to institute changes that improve upon our initial work and to monitor outcomes.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

Funding

This work was supported by funding from the Harvard Initiative for Teaching and Learning Cultivation Grant, 2014, https://hilt.harvard.edu/funding-opportunities/previoulsy-awardedprojects/projects/assessing-the-impact-of-an-innovative-curriculum-at-harvard-medical-school-a-new-paradigm-for-medical-education/

Glossary

Case-based collaborative learning (CBCL): Is a novel student-centered learning approach that borrows from team-based learning principles and incorporates elements of problem-based learning and case-based learning. CBCL includes a pre-class readiness assurance process and case-based in-class activities in which students respond to focused, open-ended questions individually, discuss their answers in small groups, and then reach consensus in a larger group facilitated by a faculty tutor.

Notes on contributors

Amy M. Sullivan, EdD, is an associate professor of medicine, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA.

Edward Krupat, PhD, is an associate professor of medicine, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA.

Jules L. Dienstag, MD, is professor of medicine, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA.

Jakob I. McSparron, MD, is an associate professor of medicine, University of Michigan School of Medicine, Ann Arbor, MA.

Stephen R. Pelletier, PhD, is senior project manager, Office of Educational Quality Improvement, Harvard Medical School, Boston, Massachusetts.

Sara B. Fazio, MD, is professor of medicine, Department of Medicine, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts.

Thomas J. Fleenor, MEd, is a data analyst in the Office of Educational Quality Improvement, Harvard Medical School, Boston, MA.

John L. Dalrymple, MD, is a professor of obstetrics, gynecology and reproductive biology, Department of Obstetrics and Gynecology, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts.

Edward M. Hundert, MD, is Dean for medical education and professor in residence of global health and social medicine and medical education at Harvard Medical School, Boston, Massachusetts.

Richard M. Schwartzstein, MD, is a professor of medicine, Department of Medicine, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts.

ORCID

Amy M. Sullivan http://orcid.org/0000-0003-1020-3197
Richard M. Schwartzstein http://orcid.org/0000-0002-8709-4884

References

AAMC. 2015–2016. Year two questionnaire. Washington(DC): Association of American Medical Colleges; [accessed 2021 Jul 14]. https://www.aamc.org/data-reports/students-residents/report/year-two-questionnaire-y2q.

Artino AR, Jr., La Rochelle JS, Deeze KJ, Gehlbach H. 2014. Developing questionnaires for educational research: AMEE Guide No. 87. Med Teach. 36(6):463–474.

Cacioppo JT, Petty RE, Kao CF. 2013. Need for Cognition Scale. Galway(IE): Measurement Instrument Database for the Social Sciences; [accessed 2021 Jul 13]. https://www.middss.org/content/need-cognition-scale.

Creswell JW, Plano Clark VL. 2017. Designing and conducting mixed methods research. 3rd ed. Thousand Oaks(CA): SAGE Publications.

Detry M, Ma Y. 2016. Analyzing repeated measures using mixed models. JAMA. 315(4):407–408.

Fitcher RM, White CB, Huang G, Schwartzstein R. 2010. Toward hypothesis-driven medical education research: task force report from the millennium conference 2007 on educational research. Acad Med. 85(5):821–828.

Fischel JE, Olvet DM, Iuli RJ, Lu WH, Chandran L. 2019. Curriculum reform and evolution: innovative content and processes at one US medical school. Med Teach. 41(1):99–106.

Frye AW, Hemmer PA. 2012. Program evaluation models and related theories: AMEE Guide No. 67. Med Teach. 34(5):288–299.

Gruppen LD. 2007. Improving medical education research. Teach Learn Med. 19(4):331–335.

Halbesleben JR, Demerouti E. 2005. The construct validity of an alternative measure of burnout: Investigating the English translation of the Oldenburg Burnout Inventory. Work and Stress. 19(2):208–220.

Heiman HL, O’Brien CL, Curry RH, Green MM, Baker JF, Kushner RF, Thomas JK, Corbridge TC, Corcoran JF, Hauser JM, et al. 2018. Description and early outcomes of a comprehensive curriculum redesign at the Northwestern University Feinberg School of Medicine. Acad Med. 93(4):593–599.

[HMS] Harvard Medical School, Medical Education. 2021. The Harvard-MIT program in health sciences and technology. Boston(MA): Harvard Medical School; [accessed 2021 Jul 14]. https://meded.hms.harvard.edu/health-sciences-technology.

Karp K, Abendroth CS. 2012. How we conduct ongoing programmatic evaluation of our medical education curriculum. Med Teach. 34(10):783–786.

Kirkpatrick JD, Kirkpatrick WK. 2021. An introduction to the new world Kirkpatrick model. Newman(GA): Kirkpatrick Partners, LLC; [accessed
Konopasek L, Norcini J, Krupat E. 2016. Focusing on the formative: building an assessment system aimed at student growth and development. Acad Med. 91(11):1492–1497.

Krupat E, Borges NJ, Brower RD, Haidet PM, Schroth WS, Fleenor TJ, Uijtdehaage S. 2017. The educational climate inventory: measuring students’ perceptions of the preclerkship and clerkship settings. Acad Med. 92(12):1757–1764.

Krupat E, Richards JB, Sullivan AM, Fleenor TJ, Schwartzstein RM. 2016. Assessing the effectiveness of case-based collaborative learning via randomized controlled trial. Acad Med. 91(5):723–729.

Krupat E, Yeager CM, Putnam S. 2000. Patient role orientations, doctor-patient fit, and visit satisfaction. Psychol Health. 15(5):707–719.

Litman JA, Crowson HM, Kolinski K. 2010. Validity of the interest- and deprivation-type epistemic curiosity distinction in non-students. Personal Individ Diff. 49(5):531–536.

Obeso V, Brown D, Ayer M, Barron B, Bull J, Carter T, Emery M, Gillespie C, Hormann M, Hyderi A, Lupi C, Schwartz M, Ullman M, Vasilevskis EE, Yingling S, Phillippi C, eds. 2017. Toolkits for the 13 core entrustable professional activities for entering residency. Washington (DC): Association of American Medical Colleges; [accessed 2021 Jul 10]. aamc.org/initiatives/coreepas/publicationsandpresentations.

O’Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. 2014. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med. 89(9):1245–1251.

Ricotta DN, Huang GC, Hale AJ, Freed JA, Smith CC. 2019. Mindset theory in medical education. Clin Teach. 16(2):159–161.

Ritchie J, Lewis J, Nicholls CM, Ormston R. 2013. The foundation of qualitative research. Qualitative research practice: a guide for social science students and researchers. 2nd Edition. Thousand Oaks (CA): SAGE Publications.

Royce CS, Hayes M, Schwartzstein RM. 2019. Teaching critical thinking: a case for instruction in cognitive biases to reduce diagnostic errors and improve patient safety. Acad Med. 94(2):187–194.

Schochet PZ. 2008. Technical methods report: guidelines for multiple testing in impact evaluations (NCEE 2008-4018). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Schwartzstein RM. 2017. Good teaching and good teachers: the role of medical education research. Annals ATS. 14(4):487–488.

Schwartzstein RM, Dienstag JL, King RW, Chang BS, Flanagan JG, Besche HC, Hoening MP, Miloslavsky EM, Atkins KM, Puig A, Pathways Writing Group, et al. 2020. The Harvard Medical School pathways curriculum: reimagining developmentally appropriate medical education for contemporary learners. Acad Med. 95(11):1687–1695.

Shenton AK. 2004. Strategies for ensuring trustworthiness in qualitative research projects. EFI. 22(2):63–75.

Smith SD, Dunham L, Dekhtyar M, Dinh A, Lanken PN, Moynahan KF, Stuber ML, Skochelak SE. 2016. Medical student perceptions of the learning environment: learning communities are associated with a more positive learning environment in a multi-institutional medical school study. Acad Med. 91(9):1263–1269.

Tavakol S, Dennick R, Tavakol M. 2011. Psychometric properties and confirmatory factor analysis of the Jefferson Scale of Physician Empathy. BMC Med Educ. 11:54.

Wong BM, Headrick LA. 2021. Application of continuous quality improvement to medical education. Med Educ. 55(1):72–81.

Yarbrough DB, Shula LM, Hopson RK, Caruthers FA. 2010. The program evaluation standards: a guide for evaluators and evaluation users. 3rd. ed. Thousand Oaks (CA): Corwin Press.