



Regular Article

Towards high reliability in national pathology education: Evaluating the United States and Canadian Academy of Pathology educational product



Cynthia K. Harris, MD^{b,1}, Yigu Chen, MPH, PMP^{a,1}, Kristin C. Jensen, MD^c, Jason L. Hornick, MD, PhD^d, Claire Kilfoyle, BA^e, Laura W. Lamps, MD^f, Yael K. Heher, MD MPH^{b,*}

^a Memorial Sloan Kettering Cancer Center, Weill Cornell Medicine, New York NY, USA

^b Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston MA, USA

^c Department of Pathology, Stanford University, USA

^d Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston MA, USA

^e Director of Education, United States and Canadian Academy of Pathology, USA

^f Department of Pathology, University of Michigan, USA

ABSTRACT

The United States and Canadian Academy of Pathology (USCAP) leadership undertook a high level, global review of educational product outcomes data using high reliability organization (HRO) principles: preoccupation with failure; reluctance to simplify; sensitivity to operations; commitment to resilience; and deference to expertise. HRO principles have long been applied to fields such as aviation, nuclear power, and more recently to healthcare, yet they are rarely applied to the field that underpins these—and many other—complex systems: education. While errors in education are less calamitous than in air travel or healthcare delivery, USCAP's educational products impact over 15,000 learners a year, and thus have important implications for the future practice of pathology. Here we report USCAP's experiences using HRO principles to evaluate our keystone educational product, the “USCAP Short Course.” Following this novel method of data review, USCAP leadership was able to better understand diverse learner needs based on practice venue, training level, and course topic. Unexpected lessons included the identification of specifically challenging educational topics, such as molecular pathology, and a need to focus more resources on emerging fields such as quality and patient safety. The results allow USCAP to assess educational product performance using HRO tools, and provide strong data-driven decision support for future national pathology education strategy.

Keywords: High reliability organization, HRO, Quality improvement, Education, USCAP, Quality and safety

Introduction

In 1906, Drs. Maude Abbott, William Osler, and others founded the United States and Canadian Academy of Pathology (USCAP)—then called the International Association of Medical Museums—to promote pathology education and scholarship.¹ Since that time, USCAP's educational reach has grown, with the organization now counting over 8000 pathologists as members.² USCAP members span the professional spectrum from pathology trainees to established practitioners, with a wide variety of clinical practice settings represented, including academic medical centers, community-based practices, reference laboratories, and industry.

As part of its core educational mission, USCAP produces educational products that are available to both members and non-members. USCAP's educational products are accredited by the Accreditation Council for Continuing Medical Education (ACCME), making participants eligible for Continuing Medical Education (CME) credits. As part of the ACCME

accreditation requirements, USCAP surveys participants on learner change.³ However, while the ACCME, similar to laboratory regulatory accreditation agencies like The Joint Commission⁴ and the College of American Pathologists,⁵ requires data collection on specific metrics, it does not necessarily require that the organization act on this information, nor is the ACCME regulatory body resourced to follow-up on action taken. Although the USCAP Education Committee (EC) has always reviewed the evaluations of USCAP's courses, the USCAP leadership lacked a systematic approach to interpreting signals from the large quantities of data gathered. To better understand diverse participant needs and to identify potential areas for improvement, USCAP's Continuing Medical Education Subcommittee sought a method to evaluate the educational product quality.

There were many educational evaluation models to choose from. The four most common, however, are the experimental/quasi-experimental approach to evaluation, Kirkpatrick's four-level evaluation model, the Logic Model, and the Context/Input/Process/Product (CIPP) model.⁶ As

* Corresponding author. Department of Pathology WRN501, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA.

E-mail address: yheher@mgh.harvard.edu (Y.K. Heher).

¹ Co-first authors.

<https://doi.org/10.1016/j.acpath.2022.100048>

Received 1 April 2022; Received in revised form 15 June 2022; Accepted 26 June 2022, Available online xxxx

eISSN2374-2895/© 2022 The Authors. Published by Elsevier Inc. on behalf of Association of Pathology Chairs. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

the name suggests, the experimental/quasi-experimental approach isolates individual elements for evaluation, much as one would in a scientific experiment.⁷ While this model is successful in the (more) controlled environment of scientific experimentation, the complex and dynamic nature of education typically precludes isolating a single variable for study. Kirkpatrick's four-level evaluation model looks at four levels of outcomes from an educational program: (1) Reaction (learner satisfaction with the program); (2) Learning (knowledge gained from the program); (3) Behavior (changes in learner behavior after the program); and (4) Results (changes in outcomes related to the program's content).⁸ Kirkpatrick's model involves a focus on outcomes that goes beyond learner satisfaction, but it has been criticized for a narrow focus on outcomes that can prevent a full evaluation of an educational program.⁹ The Logic Model proposes a clear, linear relationship between inputs and ultimate outcomes.¹⁰ The linearity and rationality implied by the Logic Model can lead an educator to (consciously or subconsciously) ignore unexpected, paradoxical, inconsistent, or highly variable outcomes.¹¹ The CIPP model looks the following components of an education program: Context (setting, utility, and resources); Input (implementation strategies, particularly feasibility, and likely efficacy); Process (assessment of the implemented program); and Product (program outcomes).^{12,13}

Although the overarching structural elements of the CIPP model were utilized when evaluating the USCAP educational product, the committee specifically sought a lens through which to understand the evaluation of the educational product. In other words, while the CIPP model is intended to evaluate educational outcomes by gathering data like participant feedback, the USCAP committee was looking for a lens through which to *understand* that feedback. The lens was high reliability.

Originally used in the fields of aviation and nuclear power, high reliability is, fundamentally, persistent mindfulness at the level of an organization that leads to consistently excellent outcomes.^{14,15} While errors in education are less calamitous than in air travel or healthcare delivery, USCAP's educational products impacted over 16,000 learners in 2017, and thus have important implications for the future practice of pathology. There are five principles of a High Reliability Organization (HRO): (1) preoccupation with failure; (2) reluctance to simplify; (3) sensitivity to operations; (4) commitment to resilience; and (5) deference to expertise. HROs are defined by their preoccupation with failure, by the refusal to "rest on their laurels" and to instead seek out imperfections. This preoccupation with failure roots out even the smallest of potential errors rather than being complacent and content with the already success achieved. Once an error is found, an HRO-mindset demands that simplistic explanations be rejected in favor of nuanced, thoughtful, and well-researched explanations. Part of that nuanced explanation involves a sensitivity to operations, a focus on the actual current performance as related by those on the frontline. Without a commitment to resilience, an HRO's preoccupation with failure could lead to a paralytic pessimism, but instead a resilient organization views failure as an opportunity for growth and further excellence. Finally, like sensitivity to operations, the principle of deference to expertise prioritizes the voices of those with hands-on knowledge and training over those in leadership positions who are abstracted from the day-to-day processes.

Using these HRO principles, we performed root cause analyses of educational product performance variation and empowered course planners and course faculty to make changes. Here we report our organization's experiences using HRO principles to evaluate and improve our keystone educational product, the "USCAP Short Course."

Materials and methods

Study period

All short courses from the 2017 USCAP annual meeting, which took place in San Antonio, Texas, on March 4 to 10, 2017, were studied.

Short course

Short courses are the most highly attended and topically diverse primary educational product at the USCAP annual meeting, the largest annual USCAP gathering. At any given USCAP annual meeting, a broad range of typically focused organ system-based short courses are offered to cover a deep and comprehensive range of pathology sub-specialties. Short courses are typically 120 min in length, can be didactic or interactive in format, and are taught by two or more faculty members who are subject matter experts. They are traditionally offered live for a 3-year period at consecutive annual meetings. They are often recorded and offered online as enduring educational products on demand.

Evaluation survey

At the end of the short course, short course learners were emailed a link to the web-based course evaluation. In accordance with ACCME requirements, CME credit was only awarded following completion of the course evaluation. Participant feedback was obtained using a 44-question course evaluation survey designed by the EC. Of the original 44 survey questions, 29 questions were included in the data analysis. The questions were categorized into four focused sections: (i) participant demographics (Q1-2); (ii) course design and delivery (Q3-11); (iii) course outcomes (Q12-22) and; (iv) learner perception of faculty performance (Q23-29; see [Supplemental Appendix 1: Original Survey Questions](#)). Section i (Q1-2) documented the participant's current role (i.e., practicing pathologist, pathologist-in-training, other) and primary practice venue (i.e., university/medical school, community practice, independent laboratory, commercial laboratory, other). Sections ii, iii, and iv (Q3-29) consisted of statements about course design, course delivery, course effectiveness, and learner perception of faculty performance. A 5-point Likert scale was used, which ranged from strongly disagree to strongly agree, with a neutral midpoint. Participants were able to skip survey questions by selecting "not applicable." Free text space was provided for general comments on participants' overall experience (Q30).

Data analysis

Raw survey data from 58 short courses were imported from USCAP's survey website and aggregated into Excel spreadsheets. "Not applicable" entries were excluded from analysis. Likert scale data were converted into numerical scores with the assumption that the interval scale of Likert categories were equal (1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree). Descriptive statistics were used to derive percentages and frequencies for categorical data. Converted Likert scale data were expressed as median, mean, and standard deviations. Violin graphs with boxplots were generated to display data distribution. A scatter plot was used to identify clusters and explore the relationship between learner perception of faculty performance and course content evaluation data by analyzing dual metrics. The one-way analysis of variance (ANOVA) test was applied to determine whether there were any statistically significant differences between the means of converted Likert scale data among different survey questions, participant roles, or primary practice venues. A p-value less than or equal to 0.05 was considered statistically significant. Statistical analyses were performed in SAS 9.3 (SAS Institute Inc., Cary, NC) and charts were generated by Power BI (Microsoft, Redmond, Washington).

Results

Survey participant demographics (section i)

There were 3617 participants registered for 58 short courses at the 2017 USCAP annual meeting with an absolute survey response rate of 29.6%. A total of 48,161 data points were collected from 1072 unique

respondents. Of the 1072 survey respondents, the vast majority were practicing pathologists (n = 981, 91.5%) with fewer pathologists-in-training (n = 80, 7.5%), and only a small number of other professionals (n = 11, 1.0%). Survey respondents were from university/medical schools (45.9%), community practices (37.5%), independent laboratories (9%), commercial laboratories (2.9%), and other venues (4.7%) (Table 1).

Survey responses for course design, delivery, and effectiveness (sections ii and iii)

Greater than 90% of respondents either strongly agreed or agreed with survey statements in section ii (Q3-11), which asserted that the short course was well-designed, delivered, and carried out effectively. The average quantified score was 4.50 (std = 0.60) out of 5 for nine questions in this section. The differences in the average quantified score between questions were statistically significant (P = 0.03). The statement “content was current and evidence-based” achieved the highest average quantified score at 4.53 (std = 0.57), while the statement “the learner was encouraged to evaluate this session” had the lowest average score at 4.45 (std = 0.62). Violin charts for Q3-Q11 (course design and delivery) show a clear bimodal distribution with clustering towards “strongly agree” (quantified score = 5) and “agree” (quantified score = 4). Every question in this section achieved a median response of “strongly agree” (quantified score = 5). (Fig. 1).

Questions in section iii (Q12-22) were designed to assess whether courses expanded learners' medical knowledge, enhanced their diagnostic/clinical practice skills, led to improvement in their performance, and/or benefitted patients. The average quantified score was 4.35 (std = 0.70) out of 5 for all eleven questions. Out of 1067 respondents, 1023 (95.7%) strongly agreed or agreed that the short course enhanced their knowledge base, and 92.8% of the respondents would recommend the course to colleagues. However, scores varied on questions like whether courses “helped create or revise protocols, policies, and procedures”, “provided knowledge, strategies, and skills to improve efficiency”, or “provided knowledge, strategies, and skills to improve safety”. Over 12% of the total responses to these three statements were negative or neutral, which led to lower quantified average scores of less than 4.30 out of 5 (Table 2). Differences observed in average quantified score between questions in section iii were statistically significant (P < 0.001). Violin charts for Q13-Q22 (related course outcomes) showed more multimodal distribution. Small groups of “neutral” responses (quantified score = 3) were identified. Only two out of eleven questions in this section had a median response of “strongly agree”, and median response for the remainder of section questions was “agree”. (Fig. 1).

Statistically significant differences were found in average quantified scores for course design and delivery statements in section ii: course design and delivery (4.35, std = 0.70) when compared to course outcome statements in section iii: related course outcomes (4.50, std = 0.60; p < 0.001).

Table 1
Survey demographics.

	N	%
Total short courses	58	
Participant status	1072	
Practicing pathologist	981	91.5%
Pathologist-in-training	80	7.5%
Other	11	1.0%
Primary practice venue	1071	
University/Medical school	492	45.9%
Community practice	402	37.5%
Independent laboratory	96	9.0%
Commercial laboratory	31	2.9%
Other	50	4.7%

Table 3 and Fig. 2 provide a breakdown of data on quantified scores grouped by respondent role. Responses were consistent between the “pathologist-in-training” group and the “practicing pathologist” group with no statistically significant difference detected for course design and delivery questions. However, a notable gap between groups was identified in results for Q13 to Q19 (questions on whether courses would prompt changes in practice). Pathologists-in-training responded with statistically significant higher agreement scores than practicing pathologists did in statements on whether the courses “helped learner create/revise protocols, policies, and procedures” (p = 0.04), “addressed professional practice gaps in learner's knowledge base” (p = 0.04), and “learned a skill that learner needs to improve or change practice” (p = 0.05).

Table 4 and Fig. 3 provide data breakdown on quantified scores grouped by respondent primary practice venue. The overall responses from those who practiced at academic medical centers and community hospitals showed similar patterns. Respondents from independent laboratories gave the lowest agreement scores to many of the survey questions, while those from commercial labs assigned highest average scores to each statement in survey section ii and iii.

When asked whether the courses provided knowledge, strategies, and skills to improve efficiency, safety, patient outcome and satisfaction (Q20-22), all attendees from 58 short courses scored these questions lower regardless of their practice level or venue. Average quantified score for these three questions was 4.31, while other course outcome questions (Q12-19) achieved an average of 4.37. The difference was statistically significant (P < 0.001).

Survey responses for learner perception of short course faculty performance (section iv)

Table 5 summarizes data on respondents' assessment of course faculty performance. The average quantified score was 4.47 (std = 0.66) out of 5 for the seven questions in this section. The survey statement on whether the “instructor utilized non-conventional techniques (e.g. audience response system)” received the lowest score (4.16, std = 0.90). More than 95% of respondents strongly agreed or agreed that the faculty member “disclosed relevant financial relationships”, “clearly had expertise in the subject area”, “slides and other educational materials were useful, effective, and appropriate”, “maintained the audience's attention”, “was enthusiastic”, and “is an effective teacher”. Fig. 4 illustrates the distribution of responses for Q23 to Q29. Q26, which addressed non-conventional interactive methods of teaching, was an outlier result. The root cause of this lower performance is uncertain; however, may represent lack of both teacher and learner familiarity with non-conventional teaching methods such as audience response systems.

Average learner perception of faculty performance scores (Q23-29) for each short course were plotted together with average overall course evaluation scores (Q3-22) in a scatter chart to examine the relationship between the two sets of metrics and a positive correlation was identified (Fig. 5). Courses in orange dots had relatively lower quantified scores on both course content and learner perception of faculty performance metrics. By contrast, courses in green dots had higher quantified scores on both course content and learner perception of faculty performance.

A subgroup analysis on ten courses (orange dots/lower overall performance) was conducted to explore the possible contributing factors for lower scores. Six of ten were identified as new short courses that were taught for the first time at the 2017 annual meeting. Five of the ten courses covered molecular pathology related topics. As a result of this subgroup analysis, data were re-categorized according to course status (“new” and “continued”) and course topic (“molecular-related” and “other”). Average scores for questions in survey section ii, iii, and iv were further calculated for each subgroup. Results confirmed that new molecular-related courses had the lowest average score of 4.28, while scores for other subgroups such as “continued molecular related courses”, “continued other courses”, and “new other courses” were higher and more consistent (Table 6).

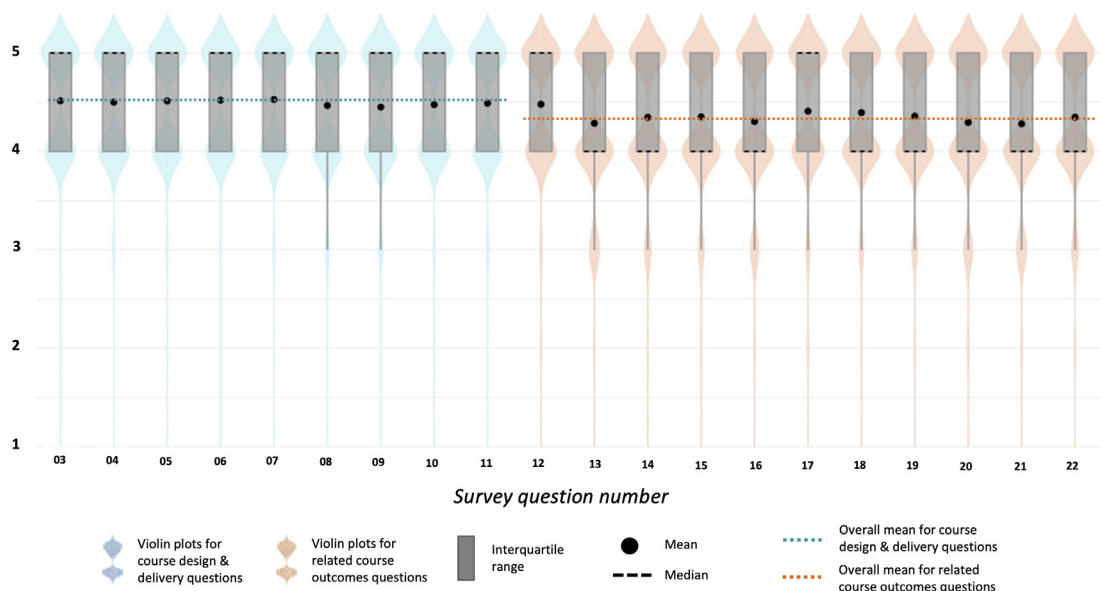


Fig. 1. Results for USCAP annual meeting 2017 short course design, delivery, and outcomes, boxplot and violin charts. The distribution of responses for Q3 to Q22 use a combination of boxplots and violin charts. Boxplots display plain data distribution by outlining the mean, median, and interquartile ranges. Violin charts offer more detail on probability density using width and length (shape area). They are useful for showing multimodality in the dataset. The shaded blue and orange areas represent the density estimate of the responses: the more answers that fall in a specific range, the larger the violin shape for that range. Solid black dots represent the response means and dashed black lines represent medians.

Table 2
Results for overall course design, delivery, and outcomes.

	Strongly Disagree = 1		Disagree = 2		Neutral = 3		Agree = 4		Strongly Agree = 5		N	Mean	Median	Std
<i>Course design and delivery</i>														
03. Content was adequately described	2	0.2%	5	0.5%	19	1.8%	457	42.8%	584	54.7%	1067	4.50	5	0.58
04. Target audience was clearly defined	1	0.1%	2	0.2%	29	2.7%	465	43.6%	570	53.4%	1067	4.50	5	0.57
05. Learning objectives were clearly stated	1	0.1%	2	0.2%	29	2.7%	452	42.4%	583	54.6%	1067	4.51	5	0.57
06. Learning objectives were appropriate	1	0.1%	4	0.4%	20	1.9%	456	42.7%	587	55.0%	1068	4.52	5	0.57
07. Content was current and evidence-based	2	0.2%	2	0.2%	25	2.3%	440	41.2%	598	56.0%	1067	4.53	5	0.57
08. Syllabus was comprehensive and well-organized	2	0.2%	6	0.6%	53	5.1%	422	40.7%	555	53.5%	1038	4.47	5	0.64
09. I was encouraged to evaluate this session	1	0.1%	3	0.3%	54	5.2%	455	43.5%	534	51.0%	1047	4.45	5	0.62
10. Pace of this session was appropriate	2	0.2%	10	0.9%	29	2.7%	464	43.5%	561	52.6%	1066	4.47	5	0.62
11. Met the stated educational objectives	1	0.1%	8	0.7%	36	3.4%	446	41.8%	577	54.0%	1068	4.49	5	0.61
<i>Course outcomes</i>														
12. Enhanced my current knowledge base	1	0.1%	13	1.2%	30	2.8%	453	42.5%	570	53.4%	1067	4.48	5	0.62
13. Helped me create/revise protocols, policies, and or procedures	3	0.3%	14	1.3%	122	11.7%	447	42.9%	457	43.8%	1043	4.29	4	0.74
14. Addressed professional "practice gaps" in my knowledge base	2	0.2%	9	0.8%	84	7.9%	490	46.2%	476	44.9%	1061	4.35	4	0.68
15. Provided me with information/knowledge to close the practice gap	2	0.2%	11	1.0%	76	7.2%	494	46.6%	476	44.9%	1059	4.35	4	0.68
16. learned a skill that I need to improve/change my practice	2	0.2%	14	1.3%	104	9.8%	480	45.3%	460	43.4%	1060	4.30	4	0.71
17. High-quality course that I would recommend to colleagues	4	0.4%	16	1.5%	56	5.3%	452	42.5%	535	50.3%	1063	4.41	5	0.70
18. Contributed to enhanced competence as a health care provider	2	0.2%	9	0.8%	58	5.5%	490	46.3%	500	47.2%	1059	4.39	4	0.65
19. I feel I am a better pathologist	2	0.2%	13	1.2%	73	6.9%	486	46.0%	483	45.7%	1057	4.36	4	0.68
20. Provided knowledge, strategies, and skills to improve efficiency	2	0.2%	19	1.8%	109	10.4%	457	43.6%	461	44.0%	1048	4.29	4	0.74
21. To improve safety in my practice	2	0.2%	19	1.8%	124	11.9%	436	41.9%	460	44.2%	1041	4.28	4	0.76
22. To improve patient outcomes and satisfaction	1	0.1%	13	1.2%	85	8.1%	474	45.0%	480	45.6%	1053	4.35	4	0.69

Discussion

Any organization seeking to improve the quality of an educational product must answer two questions. First, *who* is responsible for measuring quality? Second, *how* will the quality of the educational product be measured?

Governance of the educational product

Broadly, the USCAP's Education Committee (EC) is responsible for overseeing delivery of the USCAP educational product. The EC is an oversight body responsible for soliciting and selecting educational content and ensuring that the information presented is educationally and

Table 3
Results for course design, delivery, and outcomes group by participant role.

	Practicing pathologist			Pathologist-in-training			p-value	Other		
	Mean	Std	N	Mean	Std	N		Mean	Std	N
<i>Course design and delivery</i>										
03. Content was adequately described	4.51	0.59	975	4.58	0.52	80	0.35	4.36	0.50	11
04. Target audience was clearly defined	4.50	0.57	975	4.54	0.59	80	0.57	4.36	0.50	11
05. Learning objectives were clearly stated	4.51	0.57	975	4.51	0.60	80	0.97	4.36	0.50	11
06. Learning objectives were appropriate	4.52	0.57	976	4.59	0.52	80	0.29	4.36	0.50	11
07. Content was current and evidence-based	4.52	0.58	976	4.58	0.57	79	0.39	4.45	0.52	11
08. Syllabus was comprehensive and well-organized	4.46	0.64	946	4.53	0.59	80	0.41	4.27	0.65	11
09. I was encouraged to evaluate this session	4.45	0.61	955	4.47	0.67	80	0.72	4.36	0.50	11
10. Pace of this session was appropriate	4.47	0.62	974	4.53	0.64	80	0.47	4.36	0.50	11
11. Met the stated educational objectives	4.49	0.61	976	4.53	0.59	80	0.59	4.45	0.52	11
<i>Course outcomes</i>										
12. Enhanced my current knowledge base	4.47	0.63	975	4.54	0.57	80	0.39	4.45	0.52	11
13. Helped me create/revise protocols, policies, and or procedures	4.27	0.75	953	4.45	0.65	80	0.04	4.33	0.71	9
14. Addressed professional "practice gaps" in my knowledge base	4.33	0.68	970	4.50	0.64	80	0.04	4.40	0.70	10
15. Provided me with information/knowledge to close the practice gap	4.34	0.68	969	4.47	0.64	80	0.09	4.44	0.73	9
16. learned a skill that I need to improve/change my practice	4.29	0.72	970	4.45	0.63	80	0.05	4.44	0.73	9
17. High-quality course that I would recommend to colleagues	4.40	0.70	973	4.52	0.64	79	0.14	4.50	0.53	10
18. Contributed to enhanced competence as a health care provider	4.39	0.65	969	4.51	0.62	80	0.09	4.33	0.71	9
19. I feel I am a better pathologist	4.35	0.69	968	4.46	0.64	79	0.19	4.33	0.71	9
20. Provided knowledge, strategies and skills to improve efficiency	4.29	0.73	958	4.29	0.84	80	0.94	4.33	0.71	9
21. To improve safety in my practice	4.28	0.75	951	4.30	0.83	80	0.81	4.33	0.71	9
22. To improve patient outcomes and satisfaction	4.35	0.68	963	4.34	0.79	80	0.90	4.44	0.73	9

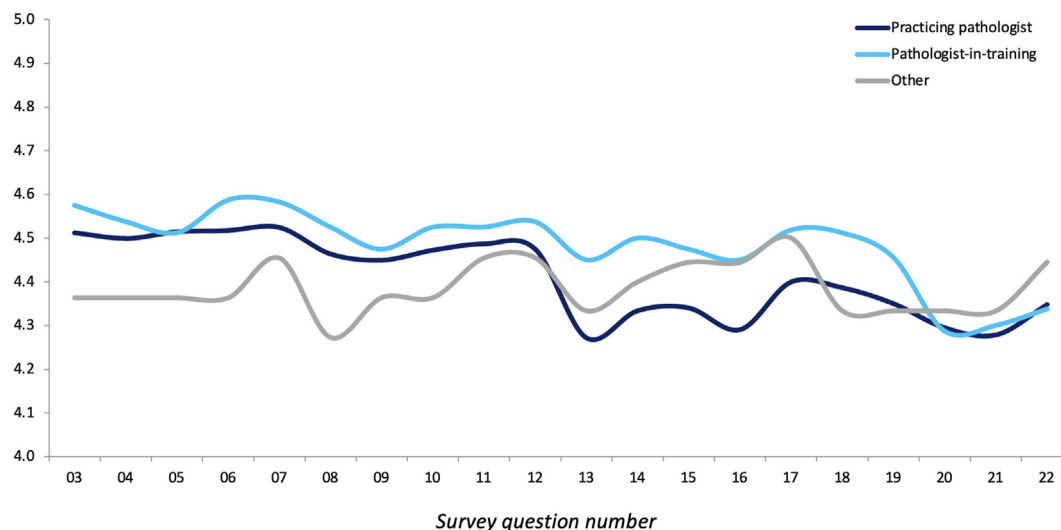


Fig. 2. Results for USCAP annual meeting 2017 short course design, delivery, and outcomes grouped by participant role: trainee versus attending pathologist and other.

scientifically relevant, high quality, evidence-based, and free of bias. The EC is a 25-person committee composed of practicing pathologists who are nominated by USCAP leadership and then elected by standing members of the EC. Typically, members have been in practice more than five years, have experience serving on USCAP subcommittees, and are published experts in their subspecialty areas.

Elected members are, while not universally, largely involved with trainee education at their respective home institutions. The Chair of the EC serves a 4-year term and is responsible for reporting the findings and activities of the EC to the USCAP Board of Directors and President (Fig. 6).

Quality assessment of the educational product

Prior to our study, the quality of the educational product was assessed in two ways: learner surveys and direct in-person auditing by EC members. As an ACCME accredited organization, USCAP collects data on learner change to assess overall program effectiveness.¹⁶ This data

collection was accomplished using online surveys filled out by learners at the conclusion of the short course. In addition to the enrolled learners, EC members also audit new courses at the live annual meetings and fill out the same course evaluation survey at the end. While the surveys provided largely quantitative data, the in-person auditing by EC members garnered more qualitative data on the overall effectiveness of the educational product, based on the institutional and specialty experience of EC members. The EC has historically reviewed the survey responses and shared them with the teaching faculty after each annual meeting; barring any major issues, it has been left up to individual teaching faculty whether or not to analyze trends in performance or make any changes. Thus, while these two methods of evaluation generated large amounts of valuable data, the outcomes were superficial and tended towards continuation rather than innovation. Like many organizations confronted with large volumes of data, the EC struggled with how best to use those data to drive broad, meaningful change, and improvement. In addition, much like laboratory accreditation, the increasing regulatory burden of the ACCME challenged even the most dedicated educational leaders in

Table 4
Results for course design, delivery, and outcomes group by primary practice venue.

	University/ Medical school			Community practice			Independent Laboratory			Commercial Laboratory			P-value	Other		
	Mean	Std	N	Mean	Std	N	Mean	Std	N	Mean	Std	N		Mean	Std	N
<i>Course design and delivery</i>																
03. Content was adequately described	4.50	0.57	488	4.56	0.59	401	4.47	0.60	96	4.65	0.49	31	0.21	4.33	0.63	49
04. Target audience was clearly defined	4.48	0.57	488	4.55	0.57	400	4.36	0.58	96	4.61	0.50	31	0.01	4.40	0.61	50
05. Learning objectives were clearly stated	4.48	0.58	487	4.56	0.57	401	4.43	0.59	96	4.65	0.49	31	0.04	4.48	0.50	50
06. Learning objectives were appropriate	4.50	0.55	488	4.56	0.57	401	4.44	0.59	96	4.65	0.49	31	0.09	4.44	0.61	50
07. Content was current and evidence-based	4.51	0.58	487	4.57	0.58	401	4.41	0.59	96	4.65	0.49	31	0.05	4.50	0.54	50
08. Syllabus was comprehensive and well-organized	4.45	0.64	469	4.50	0.64	393	4.38	0.64	95	4.61	0.50	31	0.21	4.38	0.64	48
09. I was encouraged to evaluate this session	4.50	0.63	482	4.50	0.59	390	4.33	0.66	96	4.65	0.55	31	0.02	4.39	0.49	46
10. Pace of this session was appropriate	4.48	0.59	486	4.52	0.64	401	4.31	0.65	96	4.58	0.72	31	0.03	4.36	0.56	50
11. Met the stated educational objectives	4.47	0.61	488	4.54	0.60	401	4.39	0.62	96	4.61	0.62	31	0.08	4.40	0.64	50
<i>Course outcomes</i>																
12. Enhanced my current knowledge base	4.46	0.62	487	4.53	0.62	401	4.33	0.66	96	4.61	0.50	31	0.02	4.42	0.64	50
13. Helped me create/revise protocols, policies, and or procedures	4.27	0.74	479	4.33	0.74	394	4.06	0.80	93	4.71	0.46	31	<0.001	4.23	0.77	44
14. Addressed professional "practice gaps" in my knowledge base	4.33	0.67	485	4.38	0.68	401	4.18	0.71	93	4.71	0.46	31	<0.001	4.33	0.69	49
15. Provided me with information/ knowledge to close the practice gap	4.33	0.66	485	4.39	0.68	401	4.17	0.73	93	4.71	0.46	31	<0.001	4.34	0.70	47
16. learned a skill that I need to improve/change my practice	4.29	0.71	485	4.35	0.71	401	4.08	0.77	93	4.68	0.48	31	<0.001	4.27	0.68	48
17. High-quality course that I would recommend to colleagues	4.40	0.67	487	4.46	0.70	400	4.14	0.80	94	4.65	0.61	31	<0.001	4.41	0.57	49
18. Contributed to enhanced competence as a health care provider	4.37	0.66	485	4.45	0.63	401	4.25	0.71	92	4.65	0.49	31	0.01	4.31	0.62	48
19. I feel I am a better pathologist	4.32	0.69	481	4.41	0.68	401	4.23	0.72	94	4.65	0.49	31	0.01	4.33	0.66	48
20. Provided knowledge, strategies, and skills to improve efficiency	4.27	0.75	479	4.33	0.72	399	4.26	0.77	92	4.52	0.81	31	0.23	4.18	0.75	45
21. To improve safety in my practice	4.25	0.77	476	4.31	0.74	399	4.27	0.74	88	4.45	0.93	31	0.43	4.29	0.69	45
22. To improve patient outcomes and satisfaction	4.31	0.72	481	4.41	0.65	400	4.27	0.66	92	4.58	0.67	31	0.03	4.28	0.71	47

their ability to go above and beyond simply being compliant with what was mandated. For the USCAP leadership, there was a frustrating disconnect between the high-quality, granular input gained from lengthy course surveys and its effect on specific course improvement and overall product quality. This led the EC to undertake a higher-level analysis of educational product evaluation data, through the lens of high reliability, to gain insight into current performance and identify potential targets for improvement.

Assessment of the educational product using HRO principles

Any organization can achieve a positive outcome; what sets a (HRO) apart is the *consistency* of positive outcomes.^{14,17} Achieving high reliability necessitates evaluating, analyzing, and only then improving complex systems. Despite its increasing presence in many areas of healthcare, including pathology,¹⁸ applying a high reliability framework to educational products is remarkably rare.¹⁹ Although there are five principles of

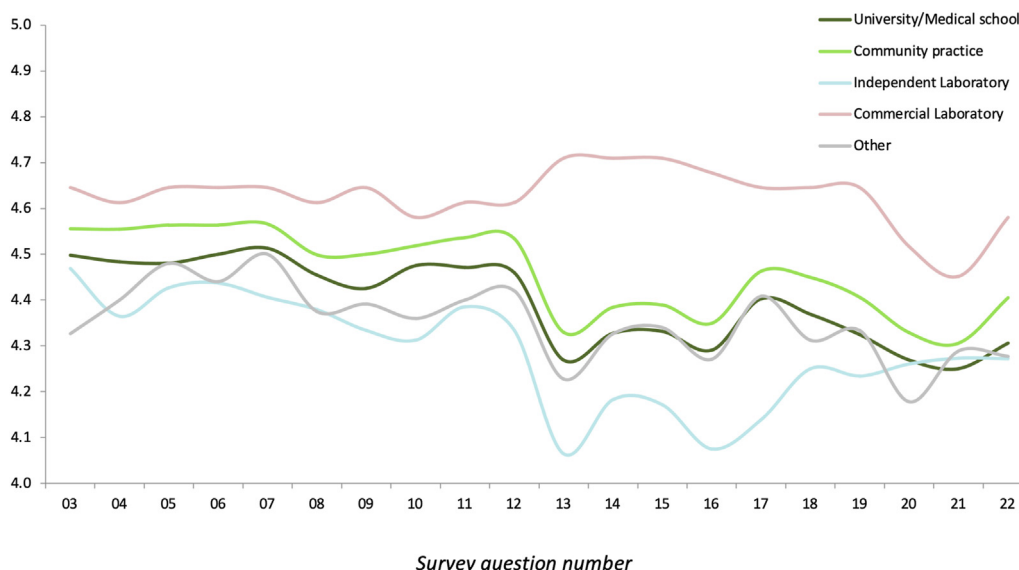


Fig. 3. Results for USCAP annual meeting 2017 short course design, delivery, and outcomes grouped by respondent primary practice venue.

Table 5
Results for learner perception of overall faculty performance.

	Strongly Disagree = 1		Disagree = 2		Neutral = 3		Agree = 4		Strongly Agree = 5		N	Mean	Median	Std
<i>Faculty Performance</i>														
23. This faculty member disclosed (before the presentation began) relevant financial	1	0.0%	3	0.1%	49	1.8%	1067	38.7%	1637	59.4%	2757	4.47	5	0.66
24. This faculty member clearly had expertise in the subject area	1	0.0%	3	0.1%	59	2.1%	1039	37.2%	1688	60.5%	2790	4.58	5	0.54
25. Slides and other educational materials were useful, effective, and appropriate	3	0.1%	12	0.4%	99	3.6%	1154	41.5%	1513	54.4%	2781	4.50	5	0.60
26. The instructor utilized non-conventional techniques (e.g. audience response systems)	18	0.7%	110	4.4%	414	16.4%	883	34.9%	1102	43.6%	2527	4.16	4	0.90
27. This faculty member maintained the audience's attention	5	0.2%	23	0.8%	102	3.7%	1124	40.4%	1528	54.9%	2782	4.49	5	0.63
28. This faculty member was enthusiastic	4	0.1%	13	0.5%	113	4.1%	1082	38.9%	1570	56.4%	2782	4.51	5	0.61
29. Overall, this faculty member is an effective teacher	7	0.3%	14	0.5%	108	3.9%	1104	39.8%	1540	55.5%	2773	4.50	5	0.62

HRO, the most fundamental for an organization seeking to improve is “preoccupation with failure.” Our educational products had received positive feedback, but rather than stop there, we analyzed specific data outliers from the broader trend of positive reception. This willingness to confront shortcomings and to seek out areas for improvement led us to analyze subgroups of learners and course topics. Our “reluctance to simplify” our findings led us to perform root cause analyses on our educational products. Although the EC has historically reviewed learners’ evaluations, the “sensitivity to operations” recommended in HRO meant the opinions of those on the frontline (learners, educators) were prioritized over the opinions of leadership. As previously discussed, the short courses are offered over a 3-year period and, while they may be modified between annual meetings, short courses are typically relatively unchanged over their 3-year run. This inherent longitudinal component

meant that the short course was the ideal focus as interventions could be undertaken prior to the completed delivery of the educational product. Following review of evaluations and root cause analysis by the EC, course planners partnered with faculty, sharing the results, and empowering them to make changes. This “commitment to resilience” ensured that the course offerings were dynamic and adaptable, rather than static. The EC is comprised of pathologists, medical education consultants, business development staff, the chief operating officer, and quality and safety analysts. This multidisciplinary committee is organized around the final HRO principle, “deference to expertise,” with the expertise of the different members prioritized over a hierarchical structure. We discuss the main lessons learned here, through the lens of HRO principles (Table 7).

1. Does the educational product lead to changes in learners' practice?

Although overall learner self-reported experience was positive, there was less agreement on whether short course experiences led to changes in practice, such as creating or revising protocols, closing practice gaps, assuring competence, or impacting efficiency and patient safety. The scores were significantly lower for this set of questions, though still on the higher end of the Likert scale.

Major barriers exist to measuring the effect of educational interventions on practice outcomes and patient safety in medicine in general, let alone in the pathology community. The first barrier is the general reluctance of medical professionals to adopt new evidence-based guidelines in a timely fashion even when they are mandated by major, validated organizations. One landmark study showed that it takes an average of 17 years for physicians to implement new guidelines broadly after they are published.^{20,21}

Another barrier lies in the very practice of pathology and the changing landscape of disease diagnosis and treatment.²² Emerging testing areas, especially in oncology and molecular pathology, are prone to rapid change and so clinical practice gaps and ambiguous recommendations are inevitable. Additionally, diagnostic testing has become complex and costly, and institutions of various types, sizes, and economic models may vary in availability of testing. Case consultation, practice subspecialization, and reference laboratory testing have all become more common. Thus, one of the results of learning more may be the awareness that instead of changing current practice to reach a new standard, it may make better sense for a subset of learners in certain practice settings (e.g., community practice or independent labs) to partner with more specialized groups such as subject matter experts or reference laboratories to provide excellent patient care. This reality makes a simple survey question around changing personal protocols, practice, and outcomes more challenging to answer positively.

2. Are there differences between types of learners?

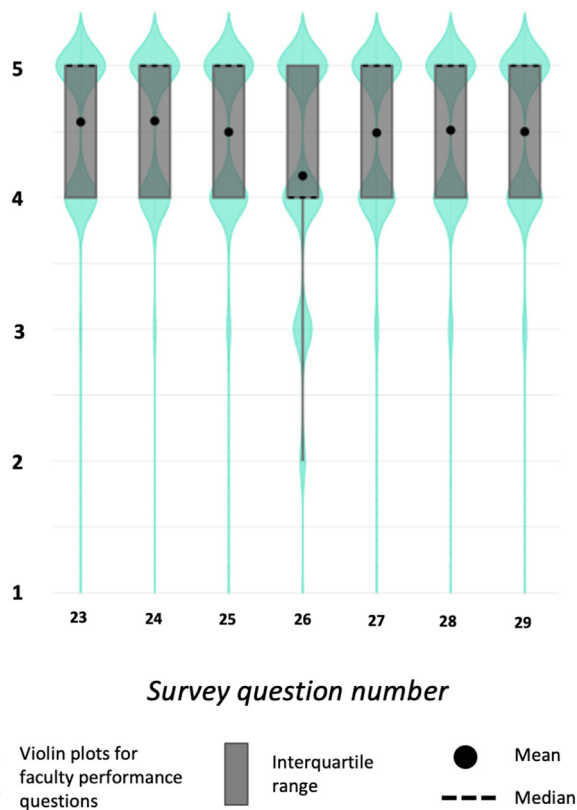


Fig. 4. Results for learner perception of faculty performance at USCAP 2017 annual meeting short courses visualized in boxplot and violin charts.

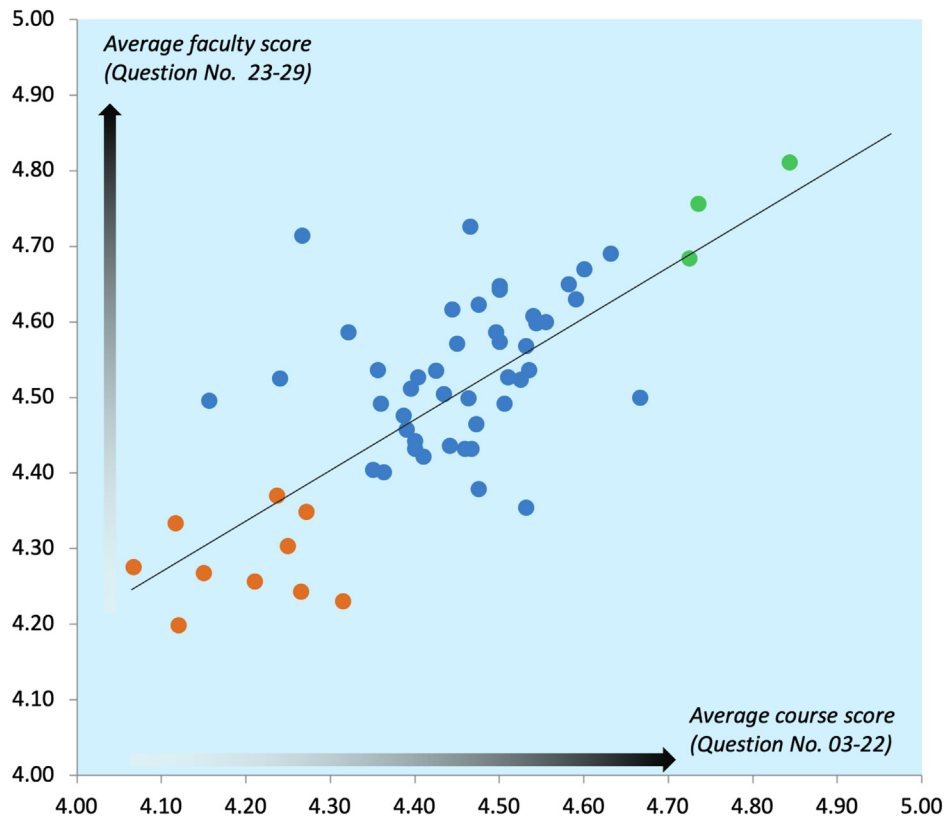


Fig. 5. Scatter chart of total average learner perception of faculty performance score and average course content score.

Table 6

Average quantified scores by course status and course topic for USCAP AM 2017 SCs.

		Course topic		
		Molecular topics	Other topics	Overall
Course status	New (1st year)	4.28	4.47	4.44
	Continued	4.47	4.44	4.45
	Overall	4.42	4.45	

Trainees versus attending pathologists

Despite general challenges in impacting learner change for pathologists in practice, one group did note a statistically significant affirmative response to questions about practice changes: pathologists-in-training. Trainees generally agreed that their learner experience would cause them to change their practice, close practice gaps, enhance their knowledge, and improve their overall performance as pathologists.

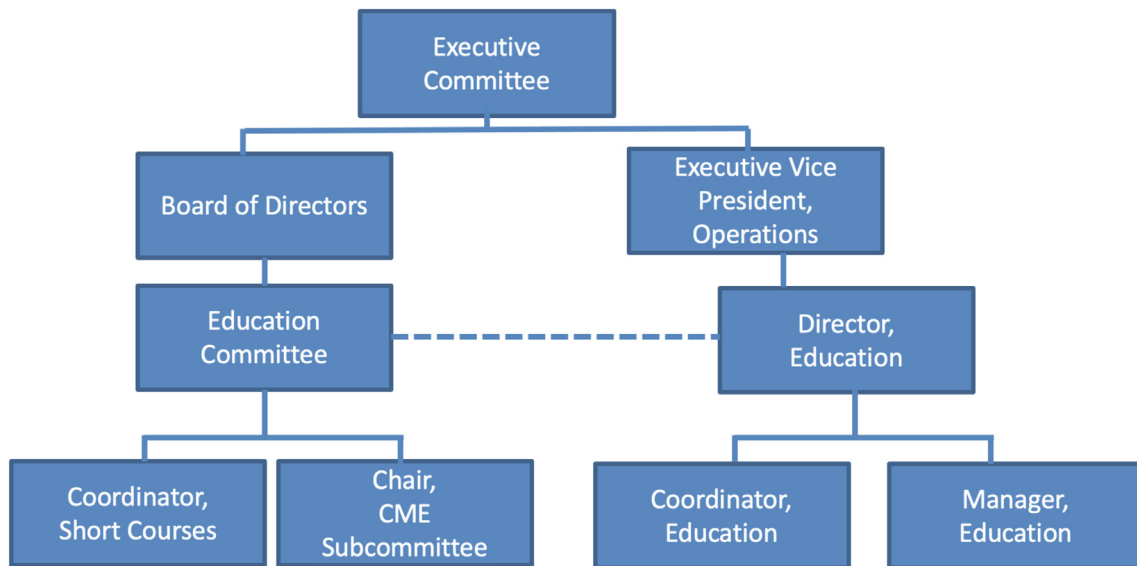


Fig. 6. USCAP Education short course educational product oversight organizational chart. The Education Committee (EC) oversees USCAP annual meeting short course content, faculty, quality, and integrity.

Table 7
High reliability organization principles as applied to USCAP educational product improvement.

HRO Principle	Definition	Conventional model	HRO-driven model	Outcome
Preoccupation with failure	Identify and analyze failures, rather than ignoring minor (or major) weaknesses	SC are generally well-received and renewed, barring major concerns	Analyze specific data outliers from broader trend of positive reception	Molecular pathology (especially new courses) and quality and safety identified as an area requiring further action
Reluctance to simplify	Embrace complexity; conduct root cause analyses; challenge long-held beliefs	A generally well received meeting means the educational product is acceptable, a poorly performing SC means a bad faculty lecturer	Analyze subgroups of learners; analyze course content quality versus teaching faculty quality, root cause analysis of “events” with educational product	Aggregate data on true root causes of poor educational product performance and act on information
Sensitivity to operations	Voices on frontline prioritized over those in leadership	Senior leaders audit portions of SC to vouch for quality	Learners directly surveyed on their experiences, feedback is analyzed	Accurate and meaningful information on the quality of the educational product from the learner perspective is collected and acted upon going forward
Commitment to resilience	Anticipating areas of trouble and recovering quickly	Feedback shared directly with teaching faculty with no discussion or action unless extenuating circumstances, SC (educational product) completes 3-yr cycle, then either extended or discontinued based on evaluations and faculty availability	Course planners partner with all faculty to make improvements annually using data gathered and root causes of challenges identified	Poorly performing educational products are continuously improved (before the following meeting) based on real, granular, actionable data directly obtained from learners
Deference to expertise	Expertise is prioritized over authority	Committee of faculty pathologists review evaluations and make recommendations for educational product and faculty members, audit courses, and make sure faculty receive comments from learners	Multidisciplinary committee comprised of pathologists, medical education consultants, business development staff, chief operating officer, quality and safety analysts, and others analyze educational product performance data, and provide actionable feedback to faculty, strategize about future courses	Future goal: continuous data feedback leads to thoughtful and deliberate designing of educational product with metrics for evaluation of outcomes on practice, continuous improvement of product delivered to learners

It is possible that practicing pathologists rated sessions lower in terms of educational value because they already consider themselves to be subject matter experts. Having deeper knowledge of the subject may make them more critical of the content and less responsive to proposed suggestions, or the information presented may confirm their current practice. For pathologists-in-training, the information being presented is likely new, therefore they are more receptive to suggested practice changes. It makes sense that learners still in training to become fully autonomous pathologists found educational content delivered from experts in the field more impactful on their practice and learning.

Academic versus community practice pathologists

Given the changing landscape of anatomic pathology testing and the trend towards referral and subspecialization, we expected to see a difference in outcomes measures between academic and community practice pathologists. This was not the case. There was no statistically significant difference between pathologist practice venue in metrics assessing outcomes or changes in practice. These data point to higher level barriers to change implementation than the basic differences between university-based practices and community practices. Barriers such as cost control, administrative hurdles, time, and a reluctance to adopt “new” practices (see item 1 above) can exist for pathologists from all backgrounds and practice settings and can be challenging to address with a single annual educational intervention.

3. Are there specific areas of weakness?

Quality and safety education

Irrespective of practice, practice venue, or course topic, learners assigned significantly lower scores to questions assessing whether courses addressed practice efficiency, patient safety, and patient outcomes. Of course, there is a lack of clear published definitions of these

relatively new concepts, particularly as they pertain to pathology. This gap likely affects the ability of learners to recognize to what extent the educational product fits into the broader framework of patient safety, outcomes, and efficiency. Recently, a new mandate from the Accreditation Council for Graduate Medical Education (ACGME) has led to refocusing on quality and patient safety in pathology training education.²³ Pathology training programs are responding to the mandate by developing and implementing quality and patient safety education nationally to meet the new Graduate Medical Education accreditation standards, and USCAP is participating in these efforts.²⁴ In addition, there has been a call to action from pathology leaders that pathologists join the patient safety movement following the publication of the most recent Institute of Medicine/National Academy of Medicine (IOM/-NAM) report “Improving Diagnosis in Medicine.”²⁵ It follows that as training programs and the pathology community adapt to adopt patient safety principles, pathologists identify knowledge and practice gaps and seek to close them. USCAP’s EC is aware of this need and has worked to address it in part by emphasizing special offerings on patient safety education. The integration of quality and safety principles into a broader set of USCAP educational products is a long-term commitment of the organization.

Molecular pathology

There was a significant difference in course performance in those courses with a topic centered on molecular pathology (Table 6). This finding is challenging to explain. Molecular pathology is a rapidly growing and changing field, where some senior practicing pathologists may feel uncomfortable and may have not received much formal training. Learners have a more difficult time with course content for which they have minimal comfortable basic knowledge. It is likely learners will show limited ability to understand and implement changes related to a topic where they are learning new, rapidly evolving, and challenging concepts for the first time, similar to the findings in the quality and safety education section.

4. How can faculty performance be separated from course performance?

Although separate survey questions existed for “faculty quality” versus “course quality,” we found that these results clustered together. For example, despite asking questions about faculty enthusiasm, expertise, and engagement, we found that these metrics followed overall course content scores. In other words, when learners rated faculty highly, they rated courses highly. This positive correlation makes intuitive sense and besides, education policy researchers have long shown that teacher quality has a strong effect on learner experience and even subsequent learner performance.^{26,27} Although the EC recognizes the value in choosing high quality faculty, these data are useful in validating that learner perception of faculty performance is closely tied to overall course experience.

Study limitations

This study looked at the survey results for one year for all learners who attended one type of educational offering: the short course. Only 29.6% of learners completed the survey. Organization-level surveys have an average response rate of 35.7%,²⁸ so our response rate is comparable to the expected response rate. However, as survey completion is tied to a learner's receiving CME credit, a selection bias may have impacted our findings, with meaningful differences between the survey respondents and the overall cohort of learners.

There are inherent challenges with studying survey evaluation data. The format is by design self-reported and highly subjective. Additionally, the results and conclusions from the survey questions are inextricable from the way questions are asked, and a poorly worded question can lead to garbled conclusions.^{29,30} Another challenge is to not overburden learners with evaluation material while still collecting as much useful data as possible. While there are some questions that must be asked as an ACCME accredited organization, USCAP worked to clarify the wording and to winnow out the valuable questions. The authors did not attempt to analyze non-numerical data (i.e. survey comments) for the purposes of this study. Although narrative feedback was utilized for individual course director planning and EC planning, it was challenging to aggregate into meaningful data for this broad, high-level study.

One of the most important ACCME metrics is the intention to change clinical practice following an educational activity. This metric is challenging to evaluate. For example, although learners may report that they gained knowledge or made practice changes, the authors were not actually able to measure the efficacy of educational interventions using validated knowledge or practice assessment tools. One future possibility would be to use pre- and post-test data for courses where this data is available and study changes in test performance as these data relate to subjective reports on educational experience. Although we did observe that it was difficult for pathologists to report a change in practice following course attendance, it may be that the time interval between when responses were measured was too short to observe an effect on practice. It may also be that a single 3-h course is, by nature, insufficient to effect practice changes, but repeated exposure to similar information over a few cycles may be sufficient.

Finally, it is important to note that this study looked at the learners' perception of course and faculty efficacy, as a surrogate marker of the actual course and faculty performance. This is an imperfect surrogate at best, as some studies have found that there is no correlation between the perception of learning and the actual learning that takes place.³¹ We did not measure learners' pre and post course knowledge, only their perception of acquired knowledge and course and faculty efficacy. This is a qualitative and secondary measurement, based on the available learner survey data. A possible future direction for USCAP educational product improvement is implementation of pre and post short course tests, which would allow for a more quantitative and primary measurement of course and faculty efficacy.

Conclusions and future directions

By analyzing USCAP's educational product through the lens of HRO principles, we were able to mindfully sift through a large data set to identify concrete lessons, hidden beneath an otherwise satisfactory performance. While learners were generally pleased with the quality of the educational product, ratings varied based on participant training level and practice venue; thus, there may be an opportunity to further customize short courses to meet the specific needs of different types of learners. A trend across all levels of training and practice venues was the identified need to integrate more patient safety and quality educational components into USCAP CME activities. The results also highlighted a few key areas for improvement including providing increased support for courses with challenging content such as molecular pathology. The EC will seek to understand these opportunities for improvement and work to further grow its offerings in these emerging areas of quality and safety and molecular pathology.

Following the in-depth review of the evaluation survey data, USCAP's EC assembled an annual meeting focus group composed of pathologist learners of diverse backgrounds, levels of training, and practice venues. These learners are being followed and studied more closely over time, with quarterly surveys and annual in-person bidirectional information exchange on educational experience, opportunities for improvement, and resources needed. Focus group data together with CME evaluation data should continue to provide useful feedback on the educational gaps and needs of pathologist learners. USCAP's EC and its leadership aim to use these data to optimize planning of future educational activities and programs to better serve pathologists, their clinical partners, and patients.

For other educational bodies interested in utilizing HRO principles to make educational improvements, we acknowledge that the level of data analysis involvement in this endeavor is daunting and even prohibitive; however, we believe that these principles can be applied in a scaled way that is appropriate to organizations of all sizes and resource-levels. As already stated above, the first two questions that must be answered are who is responsible for measuring quality and how will quality be measured. Once the team has been assembled and the mode(s) of evaluation determined (direct observation, participant surveys, pre and post examinations, comparison of outcomes with similar programs, etc.), the team can then analyze that data looking for areas of underperformance. The goal need not be about finding statistical significance, rather it can be about looking for any easily identifiable outliers. If a Likert scale is used, it is worth noting that the Likert scale tends to give very tight results with most respondents giving a rating between a 3 and a 5. This means that while a 4.3 and a 4.7 may seem similar at first glance, the difference between those ratings is actually a meaningful difference (regardless of if it is statistically significant or not). Once any outliers and areas of underperformance are identified, we recommend picking 1–3 of them to focus on improving. Be thoughtful about the planned intervention(s), implement the intervention(s), and then the following year look at the new set of data. This recommendation takes its cue from the iterative model for improvement known as the plan-do-study-act (PDSA) cycle. As the name suggests, the PDSA cycle involves identifying a goal (“plan”) (in this case, improving the educational outlier or underperformer), enacting a change (“do”) (implementing the agreed upon intervention), observing the result (“study”) (gathering the new crop of pertinent data), and then acting on what is learned (“act”) (continuing the intervention or changing the intervention).^{32,33} Finally, while we did not discuss the qualitative survey comments in this article, they are an invaluable source of feedback and ideas for directions of improvement. An educational organization looking to improve quality would be well-served by spending considerable time looking through learners' comments. By uniting HRO principles and a PDSA cycle, educational organizations can make powerful and meaningful improvements to their educational products in a short amount of time.

Funding

This study was supported by funding from the United States and Canadian Academy of Pathology.

Declaration of competing interest

The authors have no relevant conflicts of interest to disclose.

Acknowledgments

The authors would like to thank the 2018-9 USCAP Education Strategy Committee members and former Executive Vice President Dr. David Kaminsky for their initial review of the data and manuscript as well as the members of the 2018-9 USCAP Education Committee and Continuing Medical Education Subcommittee for their support of the project.

Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.acpath.2022.100048>.

References

1. The international association of medical Museums. *Science*. 1908;28(728):832.
2. USCAP history. Accessed March 15, 2022. <https://uscap.org/history/#:~:text=Our%20current%20membership%20is%20greater,conveys%20membership%20in%20the%20IAP>.
3. ACGME accreditation requirements. Accessed March 15, 2022. https://www.acgme.org/sites/default/files/2021-12/626_20211221_Accreditation_Requirements.pdf.
4. The Joint commission standards for laboratory accreditation. Accessed March 12, 2022. In: https://www.jointcommission.org/accreditation-and-certification/certification/learn-the-process/?utm_content=dm-c-135.
5. College of American pathologists laboratory accreditation. Accessed March 12, 2022. <https://elss.cap.org/elss/ShowProperty?nodePath=/UCMCON/Contribution%20Folders/WebApplications/shared-assets/lap-standards.pdf>.
6. Frye AW, Hemmer PA. Program evaluation models and related theories: AMEE guide no. 67. *Med Teach*. 2012;34(5):e288–e299.
7. Stufflebeam DL, Coryn CLS. Experimental and quasi-experimental design evaluations. *Eval Theor Models Appl*. 2014;249–290. Jossey-Bass.
8. Kirkpatrick D. Revisiting Kirkpatrick's four-level model. *Train Dev*. 1996;1:54–59.
9. Bates R. A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence. *Eval Progr Plann*. 2004;27:341–347.
10. Frechtling JA. Evaluation and logic modeling. *Logic Model Methods Prog Eval*. 2007: 1–11. Wiley.
11. Patton MQ. Developmental evaluation defined and positioned. *Develop Eval: Appl Complex Concepts Enhance Innovat Use*. 2011:1–27. Guilford Press.
12. Kellaghan T, Stufflebeam DL. The CIPP model for evaluation. *Int Handb Educ Eval*. 2003:31–62. Kluwer Academic.
13. Stufflebeam DL, Shinkfield A. Daniel Stufflebeam's CIPP model for evaluation: an improvement and accountability-oriented approach. *Eval Theor Models Appl*. 2007: 309–340. Jossey-Bass.
14. Chassin MR, Loeb JM. High-reliability health care: getting there from here. *Milbank Q*. 2013;91(3):459–490.
15. Roberson DW, Kirsh ER. Systems science: a primer on high reliability. *Otolaryngol Clin*. 2019;52(1):1–9.
16. Accreditation Council for Continuing Medical Education. The ACCME accreditation requirements. Accessed November 6, 2019. https://www.acgme.org/sites/default/files/2019-10/626_20191004_Accreditation_Requirements_Document.pdf.
17. Riley W. High reliability and implications for nursing leaders. *J Nurs Manag*. 2009; 17(2):238–246.
18. Heber YK, Chen Y, Pyatibrat S, Yoon E, Goldsmith JD, Sands KE. Achieving high reliability in histology: an improvement series to reduce errors. *Am J Clin Pathol*. 2016;146(5):554–560.
19. Stringfield S. Attempting to enhance students' learning through innovative programs: the case for schools evolving into high reliability organizations. *Sch Effect Sch Improv*. 1994;6(1):67–96.
20. Balas EA, Boren SA. Managing clinical knowledge for health care improvement. *Yearb Med Inform*. 2000;1:65–70.
21. Kanter MH, Schottinger J, Whittaker J. A model for implementing evidence-based practices more quickly. *NEJM Catal*; 2017. Accessed August 1, 2022 <https://catalyst.nejm.org/implementing-evidence-based-practices-quickly/>.
22. Balogh EP, Miller BT, Ball JR. *Improving Diagnosis in Health Care*. Committee on Diagnostic Error in Health Care; Board on Health Care Services; Board of Health Services; Institute of Medicine; The National Academies of Sciences, Engineering, and Medicine. National Academies Press; 2015.
23. The Accreditation Council for Graduate Medical Education. Common program requirements section VI table of implementation dates. Accessed November 6, 2019. <https://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/2017CPRSectionVIImplementationTable.pdf>.
24. The association of pathology chairs. Newsletter: from paths to progress. Accessed October 21, 2019. <https://www.apcprods.org/assets/newsletters/APC.Newsletter.v07-i01.pdf>.
25. Laposata M, Cohen MB. It's our turn: implications for pathology from the Institute of medicine's report on diagnostic error. *Arch Pathol Lab Med*. 2016;140(6): 505–507.
26. Nye B, Konstantopoulos S, Hedges LV. How large are teacher effects? *Educ Eval Pol Anal*. 2004;26(3):237–257.
27. Rockoff J. The impact of individual teachers on student achievement: evidence from panel data. *Am Econ Rev*. 2004;94(2):247–252.
28. Baruch Y, Holtom BC. Survey response rate levels and trends in organizational research. *Hum Relat*. 2008;61:1139–1160.
29. Kelley K, Clark B, Brown V, Sitzia J. Good practice in the conduct and reporting of survey research. *Int J Qual Health Care*. 2003;15(3):261–266.
30. Artino AR, LaRochelle JS, Dezee KJ, Gehlbach H. Developing questionnaires for educational research: AMEE Guide No. 87. *Med Teach*. 2014;36(6):463–474.
31. Uttl B, White CA, Gonzalez DW. Meta-analysis of faculty's teaching effectiveness: student evaluation of teaching ratings and student learning are not related. *Stud Educ Eval*. 2017;54:22–42.
32. Rose S, Hamill R, Caruso A, Appelbaum NP. Applying the Plan-Do-Study-Act cycle in medical education to refine an antibiotics therapy active learning session. *BMC Med Educ*. 2021;21(1):459.
33. Taylor MJ, McNicholas C, Nicolay C, Darzi A, Bell D, Reed JE. Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. *BMJ Qual Saf*. 2014;23(4):290–298.