A Consensus Approach to Identify Tiered Competencies in Quality Improvement and Patient Safety

Brennan D. Kruszewski, MD
Nathan O. Spell III, MD

ABSTRACT

Background Quality improvement and patient safety (QI/PS) competencies have been proposed separately for undergraduate medical education (UME) and graduate medical education (GME). The work forms a foundation at each educational level, yet curriculum development would benefit from more specific guidance that considers the continuum of physician training.

Objective We identified a core set of QI/PS items to be taught during medical school, residency, and independent practice, with specificity to guide curriculum development at each level.

Methods A panel of 12 QI leaders and educators with backgrounds in internal medicine from 10 academic institutions participated in consensus development using a modified Delphi technique. Three rounds of anonymous surveys were conducted, followed by a teleconference and then a fourth survey round, until consensus regarding the relevance of candidate items was reached. Items considered relevant were recommended for teaching at 1 of the 3 stages.

Results The panel identified 30 QI/PS items for learners. Of the 30 (80%), 24 were unanimously agreed on as relevant, while 6 of 30 (20%) had the agreement of 11 of the 12 experts and the assent of the remaining expert. Thirteen items were identified as appropriate for undergraduate medical education, 14 for graduate medical education, and 3 for the continuing professional development level.

Conclusions There was a high degree of agreement among 12 internists from geographically diverse institutions on the relevance of 30 QI/PS items identified for trainees in competency-based educational settings.

Introduction

Over the past 20 years, there has been growing interest in competency-based medical education frameworks and recognition that quality improvement and patient safety (QI/PS) skills are necessary for practicing physicians.[1–3] Yet existing guidance for educators who teach QI/PS is not always sufficient to design curricula. Graduates of different medical schools have variable preparation in QI/PS at entry into graduate medical education (GME) programs, and for physicians entering independent practice there is a lack of uniform expectations of QI/PS competency. The consequence is that educators in QI/PS cannot count on consistency of preparation at any educational level.

Current literature on the design of QI/PS curricula is fragmented. At the undergraduate medical education (UME) level, the Association of American Medical Colleges (AAMC) recommends in 1 of its entrustable professional activities that graduating medical students should "identify system failures and contribute to a culture of safety and improvement," with limited additional guidance.[4] The AAMC Physician Competency Reference Set also references QI/PS.[5] At the GME level, certifying boards mention QI/PS competencies,[6,7] but the detail provided is variable. In Canada, the CanMEDS framework suggests a longitudinal curriculum,[8] but a similar framework does not exist in the United States. Previous work has been done in the faculty development arena and at individual centers to establish QI/PS proficiency,[9] but educators could benefit from a tool that describes what competency in QI/PS looks like at the undergraduate, graduate, and continuing professional development (CPD) levels.

The goals of our study were to (1) establish a broad consensus about what constitutes QI/PS competency at the 3 different levels of the education continuum, and (2) to create a tool to assist educators in designing QI/PS curricula.

Methods

Educators at 15 academic institutions were identified through an author's (N.O.S.) involvement in the American Medical Association Accelerating Change in Medical Education Consortium and with the Quality and Patient Safety Subcommittee of the

DOI: http://dx.doi.org/10.4300/JGME-D-18-00210.1

Editor's Note: The online version of this article contains a full list of items and modifications by round.

Society for General Internal Medicine. Of those initially contacted, 4 individuals referred to another expert in QI/PS at their institution; otherwise, the contact was the same as the institutionally identified QI/PS expert. Twelve of 15 (80%) invited individuals agreed to participate, representing institutions in the Northeast, Mid-Atlantic, Southeast, Midwest, and Pacific Northwest. The remaining 3 did not respond to the invitation. Seven of 12 (58%) participants identified primarily as health systems leaders, and 5 of 12 (42%) identified as medical educators. All participants were board certified. The median number of years since graduation from medical school was 18.5 (range, 6–43 years).

What was known and gap
A longitudinal approach to teaching quality improvement and patient safety (QI/PS) is needed across the continuum of physician training.

What is new
Using a modified Delphi technique, 12 academic internists identified 30 core educational items for competency-based education in QI/PS.

Limitations
Single specialty expert panel reduces generalizability to other specialties.

Bottom line
The 30 educational items can guide the development of a longitudinal QI/PS curriculum.

Over a 5-month period (May to September 2017), we used a modified Delphi technique to reach consensus. This consisted of 3 rounds of surveys followed by a teleconference, and then a fourth survey.[10] We identified candidate items from a set of QI/PS program objectives developed at our institution that encompassed all levels of physician professional development. Where gaps in objectives at different educational levels were identified, we supplemented material from the CanMEDS framework and the AAMC Physician Competency Reference Set,[5,8] and the expert panel contributed additional items. To organize the surveys, we grouped items into 1 of 7 subcategories: patient safety, reliability of health care delivery, variation and value, improvement methods/systems thinking, performance measurement and public reporting, culture of safety, and care transitions. Candidate items and the modifications by round are provided as online supplemental material.

We surveyed panelists using Google Forms to test candidate items. In surveys 1 to 3, panelists rated the relevance of each item to a practicing physician of any specialty, using the scale of *not relevant*, *not relevant without major revisions*, *relevant with minor revisions*, or *highly relevant*. We provided no existing guidelines or reference materials to avoid biasing results. We dichotomized responses as *not relevant* (not relevant and not relevant without major revisions) or *relevant* (relevant with minor revisions and highly relevant). When a panelist rated an item as relevant, we asked him or her to specify whether the item should be achieved by graduation from medical school (UME), by completion of residency (GME), or through CPD during independent practice. Panelists could comment or propose changes to each item. In each round, all 12 panelists responded to the survey.

Using the method from Polit and colleagues,[11] we assigned a rating of *excellent*, *good*, *fair*, or *poor* based on K*. An *excellent* rating indicates > 95% likelihood that the level of agreement would not occur by chance alone. Items that achieved *excellent* consensus without suggestions for changes were accepted and excluded from subsequent surveys. Items that received suggestions for revision were revised and included in the next survey, while items that did not receive actionable comments and did not achieve at least *fair* consensus were dropped from further consideration. New items proposed by panelists were included in the next survey. Panelists received feedback in the form of aggregate responses for each item and proposed revisions for the next round. TABLE 1 summarizes the number of additions, revisions, and deletions by round. We preserved anonymity through the survey process until the teleconference, at which point panelists were introduced and shared opinions.

We judged relevance of candidate items using the content validity index, a tool for establishing content validity in consensus research.[12] To account for the possibility of chance agreement, K* was calculated for each individual item based on the number of experts who rated the item as relevant. For larger numbers of experts, a matrix can be constructed for the number of experts voting on each item to calculate the probability of chance agreement. The probability of chance agreement of this matrix is expressed in the following equation:

$$\left(\frac{(\text{Number of experts})!}{(\text{Number agreeing!}) * [(\text{Number of experts} - \text{Number agreeing})!]}\right) * 0.5^{[[\text{Number of experts}]]}$$

We chose percentage thresholds for assigning an item to UME, GME, or CPD. When an item received more than 50% of votes for a level, we placed it at that level. When no individual level received more than 50% of the votes, if the sum of the percentage of votes for UME and GME was more than 50%, we assigned the item to the GME level, indicating that most panelists felt the item should be achieved prior

TABLE 1
Summary of Changes at Each Round

Rounds	Added	Dropped	Sent for Revision	Accepted
Round 1	5	1	26	2
Round 2	1	2	10	16
Round 3	0	1	1	9
Teleconference	0	0	0	3
Total	6	4	37	30

to entering independent practice. The remaining items were assigned to CPD.

After 3 rounds, we held a teleconference to confirm the language of 3 candidate items, as well as the most appropriate level at which items should be taught. Six of 12 (50%) panelists participated in the teleconference. Those who were not available to participate in the teleconference were invited to give feedback via e-mail. A fourth survey was used to confirm proposed changes in wording, acceptability of the proposed changes to the panel at large, and final assignments of the recommended level at which the candidate items should be achieved by trainees. Six of 12 (50%) panelists participated, and all 12 participants were given an opportunity to voice their support of or dissent from the final list.

The Emory University Institutional Review Board deemed this study exempt.

Results

We identified 30 items with a rating of *excellent* using our consensus method, including the level where they are appropriately achieved. TABLE 2 lists our final results. Thirteen items were identified at the UME level, 14 at the GME level, and 3 at the CPD level. For 27 items, 12 of 12 (100%) participants achieved consensus after round 3, before the teleconference. For the remaining 3 items, 6 of 6 (100%) teleconference panelists expressed their agreement. Of the panelists not participating in the teleconference, none dissented from these modifications, for a total approval of 12 of 12 (100%) for the entire list.

Of the items achieving a consensus level of *excellent*, 24 of 30 (80%) attained unanimous support. For the remaining 6 of 30 (20%), 1 expert did not agree on the item's relevance in the final rating, but in each case, the dissenting expert accepted the validity of consensus on the item. No experts chose to remain outside the consensus on any item.

Discussion

In this study, 12 QI leaders and educators from geographically diverse US academic institutions reached a high degree of consensus about key items in QI/PS, and agreed on competency at the UME, GME, and CPD levels. The detailed items provide guidance for teaching QI/PS at each of these levels.

Our research builds on previous work to provide more specific tools, practices, and beliefs underpinning successful achievement of competency in QI/PS. The final list of items closely resembles previously defined items, with added specificity. Compared with CanMEDS, we placed less emphasis on informatics and data skills and their application to practice improvement; our panel emphasized the need to choose appropriate ("actionable, timely, and adequate") metrics for QI. Compared with the milestones, we list additional tools that may be useful for practicing physicians. For example, the orthopedics milestones noted the need for proficiency with an operative checklist; our panel added cognitive aids like structured communication tools and care paths, as well as checklists.[5] These milestones call for trainees to "apply common principles and techniques of quality improvement to improve care for a panel of patients,"[6] where our panelists delineated examples of those common principles and techniques, such as failure mode and effects analysis as well as the plan-do-study-act cycle.

The difference in specificity reflects 1 goal of our study—to provide information for educators designing curricula to teach QI/PS, as it provides the level of detail necessary to fill gaps in existing educational objectives. The difference in placing the emphasis on selection of appropriate metrics versus use of existing data reflects our panel's expertise and experience as QI leaders, and that the choice of appropriate metrics is integral to the success or failure of QI initiatives. In addition, the diversity of institutions represented contributed to the degree of specificity attained; when an expert contributed a revision based on his or her experience, it could be validated and refined with the rest of the panel.

The major limitation of the study was the panel makeup. Although the experience was varied in terms of geography and years since graduation, all panelists were internists working at large academic centers, reducing generalizability to other settings. In addition, only 6 participants could attend the teleconference meeting, which may bias the results in favor of those who attended. This bias was addressed by summarizing the comments of the meeting and distributing to all of the panelists, and ultimately, there was unanimous acceptance of the changes made during the meeting. Finally, consensus methods are limited in that they provide only an expert opinion level of evidence on a particular topic.

TABLE 2
Consensus Items for Competency-Based Education in the Field of Quality Improvement and Patient Safety (QI/PS)

Items	K*	Pc	Level	Expert Agreement
Culture of safety				
Promote a culture of safety by encouraging open and safe discussion of error.	1.00	<.01	UME	12/12
Report patient safety events using locally appropriate pathways.	1.00	<.01	UME	12/12
Participate in disclosure of medical errors and patient safety events to patients and families as appropriate.	1.00	<.01	GME	12/12
Respond to errors in ways consistent with "Just Culture" principles.	1.00	<.01	GME	12/12
Identify resources to support the need for personal wellness and support of "second victims" after a patient safety event.	1.00	<.01	GME	12/12
Patient safety				
Define the following terms: near miss, error, sentinel events.	1.00	<.01	UME	12/12
Can describe the Swiss Cheese model of error, including latent and active errors.	0.92	<.01	UME	11/12
Can describe hierarchy of effective interventions to prevent errors (low-level interventions, such as double checking, versus intermediate interventions, such as checklists, versus strong interventions, such as modification of environment or forcing function).	0.92	<.01	UME	11/12
Treat near misses and harmful patient safety events as opportunities for improvement.	1.00	<.01	GME	12/12
Can partner in interprofessional project teams to conduct an analysis of a patient safety event.	1.00	<.01	GME	12/12
Reliability of health care delivery				
Can describe how system factors and processes create outcomes.	1.00	<.01	UME	12/12
Describe common types of cognitive and affective biases that can impair medical decision making.	0.92	<.01	UME	11/12
Can describe features of high-reliability organizations.	0.92	<.01	UME	11/12
Incorporate strategies to promote patient safety that address human and systems factors.	1.00	<.01	GME	12/12
Actively apply the principles of situational awareness (perception, understanding, and prediction) to medical practice.	1.00	<.01	GME	12/12
Use cognitive aids such as procedural checklists, structured communication tools, or care paths to enhance patient safety and the effectiveness of care.	1	<.01	GME	12/12
Design systems-based initiatives to improve the safety and reliability of care.	0.92	<.01	CPD	11/12
Care transitions				
Describe common transitions in health care where safety gaps may occur.	1.00	<.01	UME	12/12
Demonstrate effective use of evidence-based strategies to perform a patient handoff or transfer of care.	1.00	<.01	UME	12/12
Analyze local handoff practices and contribute to process improvements to enhance the safety and effectiveness of care.	1.00	<.01	GME	12/12
Improvement methods/systems thinking				
Describe the uses of common tools for quality improvement (root cause analysis, FMEA, PDSA cycles, etc).	1.00	<.01	UME	12/12
Incorporate routine self-assessment and team assessment and performance improvement into practice.	1.00	<.01	GME	12/12
Can partner in interprofessional project teams to improve quality and safety using recognized improvement methods.	1.00	<.01	GME	12/12
Variation and value				
Identify and eliminate unnecessary care in one's own testing and treatment of patients.	1.00	<.01	GME	12/12
Promote a culture of continuous improvement by applying quality improvement methods to optimize value and decrease variability in group or system practice.	1.00	<.01	CPD	12/12

TABLE 2
Consensus Items for Competency-Based Education in the Field of QI/PS (continued)

	K*	Pc	Level	Expert Agreement
Performance measurement and transparency				
Report errors and barriers to effective and safe care using incident reporting systems.	1.00	<.01	UME	12/12
Explain the value of public transparency.	0.92	<.01	UME	11/12
Identify metrics that are actionable, timely, and adequate for improvement.	1.00	<.01	GME	12/12
Effectively employ metrics to improve quality and patient safety.	1.00	<.01	GME	12/12
Critically appraise the validity of common performance measures relevant to one's own specialty.	1.00	<.01	CPD	12/12

Abbreviations: Pc, probability of chance agreement; UME, undergraduate medical education; GME, graduate medical education; CPD, continuing professional development; FMEA, failure modes and effects analysis; PDSA, plan-do-study-act.

A next step in research would be to provide additional validity evidence for these findings with a large multispecialty study to create a blueprint for QI/PS teaching across the education continuum and across different settings. Ideally, this initiative would include representation from other specialties, and experts across the continuum of education.

Conclusion

Using a modified Delphi consensus technique, a group of 12 academic internists identified 30 core educational items for competency-based education in the field of QI/PS to guide longitudinal curriculum development in this area.

References

1. Swing SR. The ACGME outcome project: retrospective and prospective. *Med Teach*. 2007;29(7):648–654.
2. Batalden P, Leach D, Swing S, et al. General competencies and accreditation in graduate medical education. *Health Aff (Millwood)*. 2002;21(5):103–111.
3. Nasca TJ, Philibert I, Brigham T, et al. The next GME accreditation system—rationale and benefits. *N Engl J Med*. 2012;366(11):1051–1056.
4. Association of American Medical Colleges. *Core Entrustable Professional Activities for Entering Residency: Curriculum Developer's Guide*. Washington, DC: AAMC; 2014.
5. Englander R, Cameron T, Ballard AJ, et al. Toward a common taxonomy of competency domains for the health professions and competencies for physicians. *Acad Med*. 2013;88(8):1088–1094.
6. Stern PJ, Albanese S, Bostrom M, et al. Orthopaedic surgery milestones. *J Grad Med Educ*. 2013;5(1):36–58.
7. Iobst W, Aagaard E, Bazari H, et al. Internal medicine milestones. *J Grad Med Educ*. 2013;5(1):14–23.
8. Frank JR, Snell L, Sherbino J, eds. *CanMEDS 2015 Physician Competency Framework*. Ottawa, Canada: Royal College of Physicians and Surgeons of Canada; 2015.
9. Nagy CJ, Zernzach RC, Jones WS, et al. Quality improvement and patient safety milestones for the continuum of physician development. *J Grad Med Educ*. 2014;6(4):793–794.
10. Humphrey-Murto, S, Varpio L, Wood TJ, et al. The use of the Delphi and other consensus group methods in medical education research: a review. *Acad Med*. 2017;92(10):1491–1498.
11. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res Nurs Health* 2007;30(4):459–467.
12. Martuza VR. *Applying Norm-Referenced and Criterion-Referenced Measurement in Education*. Boston, MA: Allyn & Bacon; 1977.

Brennan D. Kruszewski, MD, is a PGY-1 Resident, Case Western Reserve University/University Hospitals; and **Nathan O. Spell III, MD,** is Associate Dean, Education & Professional Development, and Professor, Department of Medicine, Emory University School of Medicine.

Funding: The authors report no external funding source for this study.

Conflict of interest: The authors declare they have no competing interests.

Corresponding author: Brennan D. Kruszewski, MD, University Hospitals Cleveland Medical Center, Department of Medicine, Lakeside Third Floor, 11100 Euclid Avenue, Cleveland, OH 44106, 814.392.4691, brennan.kruszewski@gmail.com

Received March 12, 2018; revision received July 28, 2018; accepted August 24, 2018.