

Will Any Road Get You There? Examining Warranted and Unwarranted Variation in Medical Education

Eric S. Holmboe, MD, and Jennifer R. Kogan, MD

Abstract

Undergraduate and graduate medical education have long embraced uniqueness and variability in curricular and assessment approaches. Some of this variability is justified (warranted or necessary variation), but a substantial portion represents unwarranted variation. A primary tenet of outcomes-based medical education is ensuring that all learners acquire essential competencies to be publicly accountable to meet societal needs. Unwarranted variation in curricular and assessment practices contributes to suboptimal and variable educational outcomes and, by extension, risks graduates delivering suboptimal health care quality. Medical

education can use lessons from the decades of study on unwarranted variation in health care as part of efforts to continuously improve the quality of training programs. To accomplish this, medical educators will first need to recognize the difference between warranted and unwarranted variation in both clinical care and educational practices. Addressing unwarranted variation will require cooperation and collaboration between multiple levels of the health care and educational systems using a quality improvement mindset. These efforts at improvement should acknowledge that some aspects of variability are not scientifically informed

and do not support desired outcomes or societal needs. This perspective examines the correlates of unwarranted variation of clinical care in medical education and the need to address the interdependency of unwarranted variation occurring between clinical and educational practices. The authors explore the challenges of variation across multiple levels: community, institution, program, and individual faculty members. The article concludes with recommendations to improve medical education by embracing the principles of continuous quality improvement to reduce the harmful effect of unwarranted variation.

Medical education, both undergraduate and graduate, has long embraced uniqueness and variability in curricular and assessment approaches. The rationale is multifactorial: medical schools, residencies, and fellowships possess different strengths and weaknesses, promote differences in their missions, and are located in different geographic regions and communities with different needs. Changing institutional financial incentives and

pressures are also important emerging factors that contribute to variation. However, the primary goal of medical education is to ensure all learners acquire the essential competencies (i.e., educational outcomes) necessary to meet societal needs (i.e., health and health care outcomes). Outcomes should drive educational processes and design. Some variability in clinical and educational processes is justified, particularly the adaptive variability (i.e., warranted or necessary variation) that enables meeting patient and community needs. Variability that fails to meet patient and community needs is considered unwarranted variation.

Unwarranted variation in curricular and assessment practices ultimately contributes to suboptimal educational outcomes, widens the fragmentation across the medical education continuum, and by extension contributes to suboptimal health care outcomes delivered by graduates.¹⁻³ Unwarranted variation occurs at multiple levels within the health care and educational systems. Learners are nested within training programs (medical schools, residencies, and fellowships), which are nested within institutions that are

nested within communities. These nested relationships are interdependent, and this interdependence impacts clinical and educational outcomes with potentially profound effects.

In this article, we examine the effects of unwarranted variation in health care, its correlates present in undergraduate medical education (UME) and graduate medical education (GME), and the need to address the interdependency of unwarranted variation occurring between health care and educational practices. We also explore the challenges of variation across the multiple levels and conclude with recommendations to improve medical education by embracing the principles of continuous quality improvement (CQI) and reducing the harmful effect of unwarranted variation.

Unwarranted Variation in Patient Care: Definitions

Unwarranted variation has been a persistent and pernicious quality and safety problem in health care delivery. Landmark studies by Wennberg and colleagues, starting in the 1970s, found that much of the variation they documented in care delivered to patients

Please see the end of this article for information about the authors.

Correspondence should be addressed to Eric S. Holmboe, Accreditation Council for Graduate Medical Education, 401 N Michigan Ave., Chicago, IL 60611; telephone: (312) 755-7034; email: eholmboe@acgme.org; Twitter: @boedudley.

Copyright © 2022 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the Association of American Medical Colleges. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Acad Med. 2022;97:1128–1136.

First published online March 15, 2022
doi: 10.1097/ACM.00000000000004667

Supplemental digital content for this article is available at <http://links.lww.com/ACADMED/B248>.

across multiple conditions could not be explained by patient or community demographics or patients' underlying health status and comorbidities.⁴⁻⁶ Scientists have long noted the interdependent relationship between structure and process in health care delivery on outcomes.^{7,8} Unwarranted variations, especially in processes of care, do not contribute to better outcomes for patients. Unfortunately, unwarranted variation continues to be a problem today.⁶ Wennberg, in his book *Tracking Medicine*, highlights 3 specific categories of care subject to unwarranted variation: effective or necessary care, preference-sensitive care, and supply-sensitive care.⁴

Effective care, or science-informed care, is defined by the presence of high-quality scientific evidence that guides patient evaluation and treatment, meaning that the majority of patients should receive (e.g., vaccines) or not receive (e.g., X-rays for low back pain) this type of care.⁹ Suboptimal quality occurs when there is underuse or overuse of care.^{6,10,11} Preference-sensitive care is defined as care and interventions where more than one evidence-based option is available, or insufficient evidence exists for a preferred option, and outcomes for each option likely differ. For example, there are multiple prostate cancer screening and treatment options, and quality-of-life outcomes can vary substantially among them.¹² A major quality problem with preference-sensitive care is that the preference of the physician, rather than the patient, often dictates choice.^{4,11,12} Supply-sensitive care relates to the frequency and amount of care a patient receives, such as the frequency of office visits for a patient with diabetes.⁴ Preference- and supply-sensitive care are especially prone to misuse and overuse.⁶ Detrimental inequities in all 3 categories of care due to bias and structural racism remain persistent.^{8,13,14}

Within training environments, unwarranted variation in health care can be further amplified by substantial variability in curricular and assessment approaches across the medical education continuum. The nested relationships experienced by health professions students are critically important, and their interdependence contributes to the educational outcomes of a training program (see Figure 1). Some of the variation in both clinical and

educational approaches is justified based on local needs of patients and learners and available resources. Warranted variation ideally should contribute to optimal clinical and educational outcomes. However, training programs need to devote more attention to the impacts of unwarranted variation embedded within the nested, interdependent relationships as part of their CQI efforts for both clinical and educational processes to help achieve desired outcomes.

Correlates of Unwarranted Variation in Medical Education

Educational training programs are subject to the same categories of variation outlined by Wennberg and colleagues⁴ (see Table 1). Variation operates at all levels shown in Figure 1, and within each level, educational factors and processes can serve to mitigate or amplify problems of unwarranted variation in both clinical and educational outcomes. The desired integrative relationship of educational and clinical outcomes within the institutional and training program levels must emphasize the centrality of the patient (see Supplemental Digital Appendix 1, at <http://links.lww.com/ACADMED/B248>). Using this outcomes lens, we examine variation at each level, starting with the community, then moving to the institution, the program, and the faculty educators within the program.

Community-level variation

Ongoing analyses by groups such as the Dartmouth Atlas Project and Commonwealth Fund continue to demonstrate wide variability in health care costs, disparities, and quality across the United States.^{6,13} A logical question is whether such variation might be perpetuated during medical education training. Early research suggests the answer is probably yes. For example, Chen and colleagues found that family medicine and internal medicine residents who trained in higher-cost regions were more likely to be higher-cost providers regardless of where they practiced after graduation.¹⁵

Furthermore, the location and nature of a community often determines what patients and conditions will be encountered by learners within a program. Substantial variation in experience due to community factors affects the nature of the workforce that is supplied to meet patient care needs, particularly if graduates ultimately practice in a community different from where they trained. The impact of community-level variation on educational outcomes has not been fully elucidated and warrants further study. Additionally, more research is needed on whether and how learners become competent in a new community where they might care for different patient populations compared with where they trained, a component of practice-based learning and improvement (PBLI).

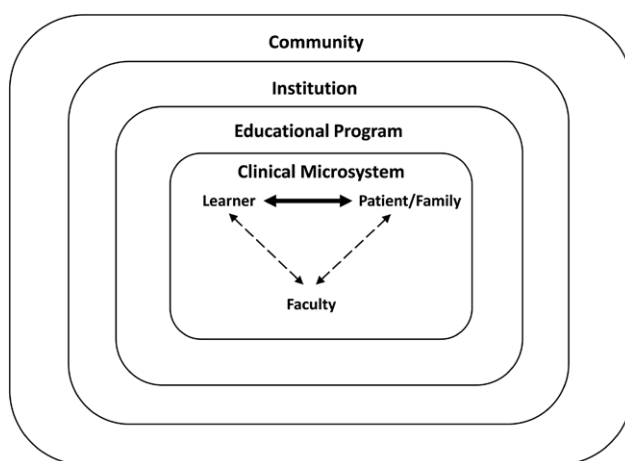


Figure 1 In the latter stages of undergraduate medical education and in graduate medical education, learners work and learn in clinical environments, most notably microsystems. The community has a major impact on who the learners will care for. Institutional and educational programmatic culture and resources will substantially shape learners' development. Learners will form bonds of variable degrees with patients (e.g., thickness of the bidirectional arrow). Faculty involvement can be highly variable with both learners and patients (strength of connection represented by the dotted arrows).

Table 1
Correlates of Unwarranted Variation in Medical Education

Wennberg category	Examples of medical education correlates
Science-informed care	<ul style="list-style-type: none"> • Variable teaching and application of effective clinical care practices • Variable teaching and assessment of evidence-based practice competency (practice-based learning and improvement) • Variable implementation of effective interprofessional team-based care (e.g., interprofessional teamwork required for care of patients with chronic illness) • Variable faculty ability in the competencies of systems-based practice and practice-based learning and improvement
Preference-sensitive care	<ul style="list-style-type: none"> • Variable teaching and assessment of effective informed and shared decision-making approaches • Variable adoption of co-production practices with patients and learners (e.g., variable engagement of patients and learners as partners) • Variable faculty ability in the competencies of interpersonal skills and communication
Supply-sensitive care	<ul style="list-style-type: none"> • Variable faculty and institutional practices around specific approaches to care (e.g., visits, procedures, lab/diagnostic testing) • Variable faculty ability in competencies of patient care, medical knowledge, systems-based practice

Institutional-level variation

Substantial variation continues to exist between training institutions that cannot be fully explained by risk adjustment or population demographics.^{6,16,17} For example, Wennberg and colleagues found substantial variation in care between training institutions with similar patient demographics in 2 separate cities.¹⁶ This institutional-level variation also appears to affect educational outcomes. In oft-cited studies, Asch and colleagues found a strong association between the rate of major obstetrical complications at a teaching hospital and the rate of major obstetrical complications among the hospitals' graduates in practice.^{18,19} This relationship was noted to persist for over 15 years. Subsequent studies examining institutional-level performance in surgical complications and costs of care found similar patterns.^{20,21}

The Accreditation Council for Graduate Medical Education (ACGME) Clinical Learning Environment Review (CLER) program has also found substantial variation in how institutions engage trainees in quality and safety processes. It can be very difficult to concomitantly improve health care and education without including learners in the process. In the most recent CLER report, less than a quarter of institutions reported including trainees in their institutional-level quality improvement and patient safety (QIPS) initiatives.²² This highlights the potentially problematic lack of integration between institutional care delivery performance and QIPS education, with significant implications for both learners and patients.

Institutional-level variation can be magnified when learners rotate at multiple hospitals and other clinical settings. Each training site has its own culture and resources that contribute to clinical and educational variability. Quality management around clinical and educational standards and practices across multiple institutions is often limited, potentially creating an amalgam of unwarranted variation within programs. This situation may contribute to suboptimal care for patients and impede professional development if medical students, residents, and fellows are not properly prepared to work in a new environment.

This shifting of learners between institutions is often justified as a mechanism to provide learners with broader experience, a reasonable educational rationale. However, standard protocols for managing the transition between institutions are lacking.²³ Lack of attention to the impact on patients and assessment of whether such curricular practices improve educational outcomes may introduce unwanted and unwarranted sources of variation. In addition, harmful and unnecessary variation (misuse, overuse, underuse) may result from misaligned financial incentives within the institution that can filter down to learners.

Program-level variation

Accreditation and certification create high-level standards around curricular content and experience, and while programs are nested within institutions,

and often more than one institution, they still have meaningful control and oversight of their curricula and assessment. Yet, the intersection of regulatory requirements with institutional needs can result in unintentionally rigid educational structures.^{24,25} For example, many medical schools and residency training programs rely heavily on a rotational training model to meet regulatory (e.g., broad coverage of curricular content) and clinical service needs. However, what happens educationally (i.e., processes) during these rotations can vary substantially from school to school and residency to residency. Learner development can vary substantially depending on when and where specific rotations are scheduled, and rotations are often scheduled independent of a learner's developmental readiness.

Additionally, the variable approach to managing transitions between rotations has serious, negative ramifications for learners and patients.^{26–29} For example, Denson and colleagues found an association of increased patient mortality in internal medicine programs when patients were discharged during intern and resident service transitions.^{27,28} Bernabeo and colleagues uncovered numerous stories of the negative impact of rotation transitions on patients, learners, and other health care professionals.²⁵ Interns and residents described developing their own "system" for these transitions in the absence of programmatic or clinical care standards. Faculty were often unaware of the impact of these rotational transitions, with some faculty arguing such transitions were simply part of trainees' development. These studies highlight that even a single source of programmatic variability, such as how transitions occur and are managed, can have profound effects—an argument strongly in support of tighter standards around rotational transitions.

Compounding this problem is the general reticence of programs to feedforward learner information to faculty from one rotation or clinical experience to another.³⁰ Yet a growing number of studies are documenting what has been known anecdotally for a very long time—learners experience different learning curves across different competencies.^{31–34} McGaghie has long argued that medical education should use a mastery-based

approach, requiring a willingness of the medical education enterprise to let go of current assumptions that dwell time, rotation-based curriculum designs, fragmented attending coverage and supervision, and single-point-in-time high-stakes assessments are sufficient to produce optimal outcomes.³³ While there are legitimate issues for medical schools regarding Family Educational Rights and Privacy Act regulations, they do not preclude feedforward activities within the medical school if proper policies and procedures are in place.³⁵ For GME, feedforward should become a standard practice to support professional development, recognizing the feedforward process must be done effectively to avoid unintended consequences.

Mastery-based learning also depends on robust, systematic assessment.³³ Unfortunately, programs across the continuum vary substantially in both their approach to assessment and the quality of their assessment activities. Twenty years after the launch of the Outcome Project, most GME training programs still struggle to assess the competencies of PBLI and systems-based practice (SBP). However, an ongoing commitment by institutions and training programs to advance PBLI and SBP is still urgently needed.³⁶ The quality and patient safety measures and robust multisource feedback tools remain substantially underused in GME programs.^{22,37,38} Clinical competency committees (CCCs), while an important addition to the assessment system, are still learning how to effectively synthesize and collate assessment data and fill assessment gaps. CCCs remain heavily dependent on faculty assessments that are one of the largest sources of unwarranted variation.³⁹

Faculty-level variation

Differences in faculty clinical and educational competence are underappreciated sources of unwarranted variation. For some researchers, faculty variation in teaching and assessment is viewed positively.^{40,41} For example, idiosyncrasy in assessment can be seen as a matter of 2 faculty members each emphasizing different but legitimate aspects of performance. If that idiosyncrasy aligns with effective, science-informed practice, the difference between faculty can be useful, warranted variation that can contribute to richer feedback and a fuller assessment of learner competence

(provided the learner recognizes it as such).^{40,41} If idiosyncrasy results from outdated or ineffective clinical practices, or the inappropriate execution of preference- or supply-sensitive care, the variation is unwarranted, thereby undermining the quality of the assessment of learner competence. There are significant long-term implications related to faculty as an unwarranted source of variation in assessment and supervision of clinical care.

Faculty clinical competence. Programs operate from the assumption their faculty possess all the necessary competencies to serve as clinician, teacher, assessor, and coach. Furthermore, supervising faculty competence in QIPS, clinical skills, diagnostic accuracy, direct observation, assessment, and evaluation are commonly, though tacitly, considered “competent until proven incompetent” by the public and by regulators. The same assumptions are made about institutions and programs. While the vast majority of clinician–educators are unquestionably dedicated to their role, the assumption that all faculty possess all needed competencies is deeply flawed.^{42–44} The majority of current clinician–educators trained in an era before there was full and effective training in the 6 ACGME general competencies. Furthermore, institutions have not sufficiently invested in faculty development in the newer competency domains such as QIPS, interprofessional teamwork, and systems thinking. Even for the 2 competencies with which most faculty feel comfortable, medical knowledge and patient care, studies demonstrate that faculty possess variable levels of competence, and this variability has real-world implications for learners.^{42,44} Learners often recognize that faculty have disparate approaches to patient care that can lead to discordant feedback as well as variable approaches to supervision. This situation can create substantial difficulties for learners to determine what feedback is most ideal to incorporate into their own practice.

Faculty assessment. Unwarranted variation commonly surfaces in faculty assessment of learners. It is well established that low interrater agreement using rating scales for work-based assessment can be problematic for both learner and patient.⁴⁴ Studies have shown faculty often use self as their primary frame of reference and attend to different aspects of a clinical encounter when directly observing a

learner with a patient, often missing clear errors in care.^{39,42,44} While assessments across faculty can be combined over time to provide an aggregate picture of competence, a given patient often receives care at a single point in time such as an ambulatory visit. Therefore, a correct assessment must be made in each encounter for the patient to receive quality care, even if the assessment is only based on a learners’ oral presentation to faculty.

Figure 2 highlights 2 faculty members assessing the performance of a trainee differently for the same encounter, with one faculty member missing an element of unacceptable care and another faculty “observing” performance in the best practice and acceptable zones that was not actually done by the trainee. Combining these observations becomes a mishmash of accurate and inaccurate observations and assessment that may lead to suboptimal feedback and supervision decisions. This variability also undermines the credibility of feedback and leaves to chance whether the learner incorporates the feedback that was most accurate. Finally, CCCs have no straightforward mechanism of making critical distinctions between the 2 assessments. For patients, this variation between the 2 faculty members, especially the unwarranted variation in the care they received, simply equates to suboptimal and potentially unsafe care.⁴⁴

Faculty also vary in their approach to supervision of learners, using different combinations of factors and approaches that can be exacerbated by variable approaches to assessment.^{45–47} This variability in supervisory styles creates confusion for learners and potentially puts patients at risk.^{48,49} Schumacher and colleagues found essentially little to no correlation between supervisors’ entrustment ratings of learners and the quality of care delivered to children in a pediatric emergency department.⁵⁰ These studies and others suggest a potentially harmful interaction between faculty (in)abilities in clinical competencies and assessment skills, and subsequent supervision decisions, especially when rating scales are employed as the primary means of making assessment judgments.

The Variation Cascade

While the interactions between the hierarchical levels of variation can have

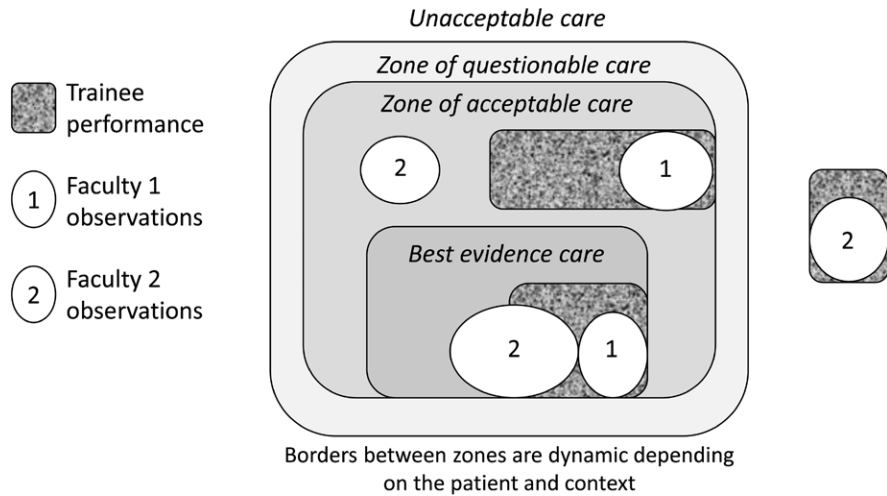


Figure 2 highlights the zones of care from a learner-patient encounter perspective, showing the implications of faculty idiosyncrasy and variable competence of the learner and faculty skill in assessment. “Quality of care” can be viewed as a set of performance zones: best evidence care, acceptable but not optimal care, questionable care, and unacceptable care. Let’s use informed decision making (IDM) as an example. Multiple studies have provided guidance on effective communication practices for IDM, such as how to set the stage for the conversation, how to explore patient preferences, use of teach back, and so forth.^{42,61} Best evidence care would include these evidence-based behaviors. Acceptable care might include the behavior of asking if the patient has questions but neglecting a teach back. Unacceptable care would be interrupting the patient so they cannot ask a question about the medical decision. The trainee’s performance in each of these zones is represented by the tiled area. In other words, the trainee performed some, but not all, of the evidenced-based elements of IDM. As we can see, the trainee also provided an element of unacceptable care in the encounter, such as dismissing a patient question or concern. Faculty 1 detected some, but not all of the best evidence care and acceptable care provided by the trainee but completely missed the component of unacceptable care. From a sampling perspective, Faculty 1 could contribute some meaningful assessment information about the trainee. From the patient perspective, however, this is a supervision “system failure” by failing to recognize and address the unacceptable care. Faculty 2 detected some, but different, behaviors of best evidence care and the unacceptable care, a better outcome for the patient compared to Faculty 1 but resulting in a very different assessment for the learner.

deleterious implications for patients, variation can also occur at multiple points during a patient care episode within teaching settings (see Figure 3). At each stage of the encounter, multiple

learner and faculty factors are present that can enhance the probability of unwarranted variation, or frank clinical errors, to the detriment of the patient. Furthermore, faculty can choose whether

to directly observe the learner at each stage. For example, in family medicine and internal medicine, faculty typically “see” patients indirectly by listening to a resident’s case presentation. This

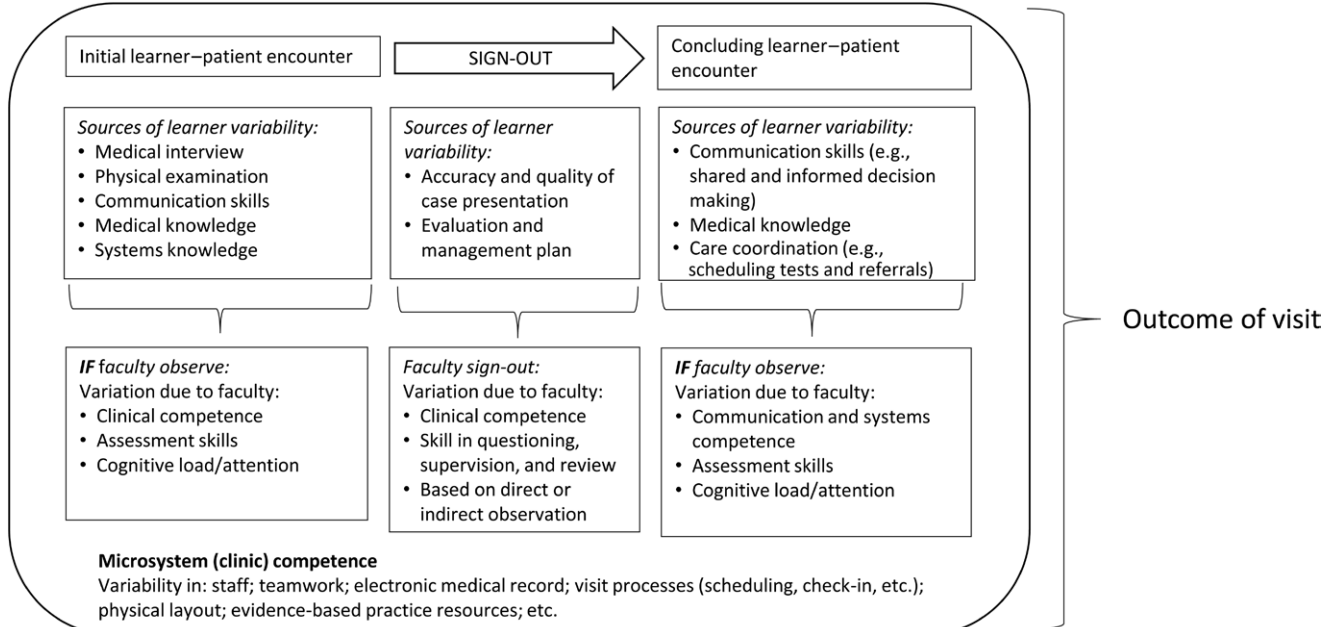


Figure 3 Variation, both warranted and unwarranted, can occur at multiple points in a single patient-learner encounter. Unwarranted variation experienced in the earlier stages of the encounter may be carried forward in the latter stages of the encounter and post-encounter, potentially and negatively affecting quality and safety for patients and learner development.

requires the sign-out process to be systematic and accurate. Unfortunately, case presentations may not accurately represent what transpired in the resident–patient encounter.^{51,52} In our faculty development work, we have heard countless stories of faculty experiencing profound discordance between what they heard from a case presentation and what they discovered when they saw the patient directly.⁵³ Singh and colleagues also found that diagnostic error in an ambulatory visit was a common reason for unexpected patient return visits within 14 days, highlighting the importance of appropriate supervision of accurate medical interview and physical examination skills for patients presenting with new signs and symptoms.⁵⁴ This is not meant to criticize faculty, but rather to highlight that learners develop and achieve competence at different rates.

Variable approaches to observation and supervision appropriately aligned with learners' developmental stage is logical, but misaligned and ineffective approaches can lead to significant deficiencies in both clinical care and education outcomes. Furthermore, failure to identify unwarranted variation or frank clinical errors negatively impacts the learner's own development, such as entrenching suboptimal skills or experiencing the emotional impact of a consequential error. It is essential that faculty work to acquire high levels of competence in both the competencies of health care and teaching and assessment, as clinical care and education are highly interdependent (see Supplemental Digital Appendix 1, at <http://links.lww.com/ACADMED/B248>).

Suggested Steps to Address Unwarranted Variation in Medical Education

Activism as a philosophy is now needed to create meaningful change in medical education. Medical education must assume the tenets of CQI and innovation to methodologically address prevailing and emerging challenges in educational programs. With this foundational purpose, medical education could be a prime mover to help achieve a transformed future state for health care.

There are several changes to UME and GME curricula and assessment

at the institutional, program, and individual learner levels that could reduce unwarranted variation while concomitantly improving health care and medical education. At the institutional level, hospitals and clinics should systematically involve learners in identifying and reducing quality gaps and unwarranted variation by targeting those measures with substantial clinical impact.⁵⁵ Learners, being frontline care givers, are ideal sources of information to detect and then address unwarranted variation as part of institutional CQI. GME programs can work with their institutional leadership to incorporate learners into QIPS initiatives or support learners in specialty-specific QIPS projects.⁵⁶

Seven curricular and assessment targets focused on programs and their faculty can and should be implemented on a wide scale to accelerate improvements (see Table 2). Implementing these recommendations will cause or exacerbate numerous tensions such as standardization versus flexibility, learner-centeredness versus patient-centeredness, and formative versus summative assessment. Navigating these tensions requires a “both/and” polarity management approach.⁵⁷ Polarity thinking requires medical educators to examine tensions as “poles” that each have benefits (“upsides”) and also disadvantages (“downsides”). Unwarranted variation, by definition, is unwanted and harmful and does not represent a viable upside choice or pole. As programs and institutions struggle with these tensions, warranted and unwarranted variation should serve as a set of lenses for making curricular and assessment choices.

Finally, a co-production mindset and approach can support CQI and further reduce unwarranted variation. Co-production in health care, as defined by Batalden and colleagues, is “the interdependent work of users and professionals to design, create, develop, deliver, assess and improve the relationships and actions that contribute to the health of individuals and populations.”⁵⁸ Co-production means that patients and learners must be critical partners with faculty, program, and institutional leadership in reducing unwarranted variation in

health care and medical education.⁵⁸ Involving learners in co-production creates additional dividends by better supporting professional development, self-regulated learning, and reflective practice.

Both medical schools and GME programs can advance CQI, co-production, and reduce unwarranted variation by adopting the following practices:

- Engaging learners as partners in curricular and educational system design;
- Individualizing learning pathways for learners where possible by incorporating effective science-informed educational practices;
- Enabling learners to “own” their own performance portfolio that includes quality and safety measures that support learning and growth;
- Optimizing opportunities for learners to develop meaningful, longitudinal relationships with patients, peers, and teachers;
- Engaging and supporting learners in guided reflection and self-assessment; and
- Engaging patients in formal teaching roles that involve direct attention to issues of warranted and unwarranted variation.⁵⁹

It will be exceedingly difficult to advance competency-based medical education if programs do not involve learners in addressing variation within the program.

Conclusions

Medical education institutions and programs have a long history of embracing uniqueness and honoring the nuances, or art, of patient care. Some of this variability is necessary, justified, and even valuable; but much is not and represents unwarranted variation that can lead to suboptimal outcomes for both patients and learners. One of the primary goals of medical education is to contribute to improving health and health care; reducing unwarranted variation in outcomes is an essential component so that learners are equipped with the competencies essential to practice and

Table 2

Recommendations to Accelerate Improvements in Medical Education by Reducing Unwarranted Variation

Component	Rationale for recommendation
Curriculum	
<ul style="list-style-type: none"> Examine current curricular design (e.g., rotations) for learner developmental appropriateness. Educational systems should track learner progress using evidence-based outcomes and use this information to make appropriate adjustments to learners' rotational assignments. This will require UME and GME systems to invest resources to create greater capacity and flexibility to accommodate different rotational sequences and rates of learner growth. 	<p>Progressive sequencing based on developmental readiness is an essential component of CBME. Programs should ask if the process for creating and then assigning learners to rotations in a specific order is appropriate to support learner development or if rotations are scheduled based on convenience and service needs. At a minimum, programs should employ assessment practices that help the program and faculty understand where the learner is developmentally even if substantial redesign is not feasible.</p>
<ul style="list-style-type: none"> Use mastery-based principles in curricular design (in support of recommendation #1). 	<p>Mastery-based education means excellence is expected and achievable by all learners who are able, motivated, and work hard. Mastery-based education implies little to no variation in measured outcomes. Mastery learning depends on learning a sequence of less complex components. If learners receive the optimal quality of instruction and learning time required, most learners should attain mastery.³³</p>
<ul style="list-style-type: none"> Create meaningful and psychologically safe feedback and feedforward mechanisms at the institutional, program, faculty, and learner levels. This may require a revision of existing policies; learners should be a participant in this process. 	<p>Feedback and coaching are essential for professional development. However, feedforward, also known as a "warm handoff," is rarely performed because of legitimate concerns about FERPA and bias. However, lack of effective educational handoffs can substantially hinder both teaching and assessment.</p>
<ul style="list-style-type: none"> Engage and empower learners through co-production around curricular design and execution to ensure the educational experience is optimized for professional developmental outcomes. 	<p>Co-producing curricular design with learners can help to identify opportunities for educational and clinical care improvement and ameliorate deficiencies and unwarranted variation in clinical and other educational experiences.⁵⁹</p>
<ul style="list-style-type: none"> Support robust faculty development in general competencies. There should be a mandate at the institutional level to support faculty development, and this should be seen as a core element of institutional CQI efforts. 	<p>Variable levels of clinical and pedagogical competence among faculty are a major contributor to unwarranted variation. Faculty development is not a luxury; institutions need to reframe their perspective to support faculty in their own professional development.</p>
<ul style="list-style-type: none"> Incorporate institutional- and program-level clinical and educational quality and safety measures into curriculum. These measures can guide both programmatic improvement and individual learning plans. 	<p>It is very difficult to identify and address unwarranted variation if it is not measured. QIPS measures exist for all 3 types of care outlined by Wennberg.⁴ Robust performance measurement is essential to improvement efforts.</p>
<ul style="list-style-type: none"> Apply evidence-based program evaluation methods longitudinally and continually. 	<p>Program evaluation is essential for continuous improvement and addressing areas of suboptimal curricular practices. This includes ensuring the learner assessment system is multifaceted and longitudinal, and processes are in place to review and act on learner- and program-level data.</p>
Assessment	
<ul style="list-style-type: none"> Implement programmatic assessment using a standard core set of assessment methods. Programs should use learner outcomes and patient outcomes as key data to inform program evaluation. Too often process measures and learner satisfaction are used as proxies for outcomes. 	<p>Programmatic assessment, using a core set of methods, is essential for assessing all the important competencies. No single assessment is sufficient to holistically judge whether an individual learner is ready for each stage of training.</p>
<ul style="list-style-type: none"> Use a developmental mindset, grounded in mastery-based learning for programmatic assessment. 	<p>Single, point-in-time assessments, especially proxy measures such as multiple-choice question tests, have limited utility. Assessment must embrace a developmental mindset.</p>
<ul style="list-style-type: none"> Create meaningful and psychologically safe feedback and feedforward mechanisms for learners. This may require a revision of existing policies; learners should be a participant in this process. 	<p>Feedback and coaching are essential for professional development. However, feedforward, also known as a "warm handoff," is rarely performed because of legitimate concerns about FERPA and bias. However, lack of effective educational handoffs can substantially hinder good assessment for learning and professional development.</p>
<ul style="list-style-type: none"> Engage and empower learners through co-production in assessment activities: learners should both seek and perform assessment. 	<p>Learners must be empowered to both seek out assessment for learning and provide input into what and how assessments are used in the training program. Assessment is something we do with learners, not just to them.</p>
<ul style="list-style-type: none"> Support robust faculty development in assessment, a complex skill that requires training. 	<p>Unwarranted variation is substantial in faculty assessments, leading to poor reliability, poor accuracy, and confusing heterogeneity and idiosyncrasy for learners and programs.</p>
<ul style="list-style-type: none"> Include group process (e.g., CCC) and evidence-based practices in assessment programs to make developmental judgments. 	<p>Group process, when done well, can reduce unwarranted variation including bias. Good evidence exists on effective group process that is not currently being fully used.</p>
<ul style="list-style-type: none"> Incorporate QIPS measures as formative, lower-stakes assessments. 	<p>QIPS measures around the 3 types of care outlined by Wennberg can support individual learner development and program improvement.</p>

Abbreviations: CBME, competency-based medical education; UME, undergraduate medical education; GME, graduate medical education; FERPA, Family Educational Rights and Privacy Act; CQI, continuous quality improvement; QIPS, quality improvement and patient safety; CCC, Clinical Competency Committee.

can function in our often fragmented health care delivery system. Further, ensuring that graduates can effectively practice CQI and reduce unwarranted variation is part of the solution for fixing an imperfect system. We must pay more attention to determining what is acceptable and unacceptable variation in approaches to patient care and medical education, and we must illuminate the interdependencies between the 2.⁶⁰ To start, confronting unwarranted variation in clinical care, poor quality of care practices, and patient safety problems is a logical target for all medical schools, residencies, and fellowships that will have direct educational benefits. Implementing science-informed educational changes and interventions can help concomitantly improve patient care and our learners' and faculty's professional development. The time is now to address and tackle the problem of unwarranted variation in medical education. Doing so could ultimately improve the educational outcomes of our learners and, most importantly, the health and outcomes of our patients and their communities.

Acknowledgments: The authors would like to thank Paul Batalden, MD, and Jung Kim, PhD, for their thoughtful review of the manuscript.

Funding/Support: None reported.

Other disclosures: E.S. Holmboe works for the Accreditation Council for Graduate Medical Education and receives royalties from Elsevier Publishing for a textbook.

Ethical approval: Reported as not applicable.

E.S. Holmboe is chief, research, milestones development, and evaluation, Accreditation Council for Graduate Medical Education, Chicago, Illinois; ORCID: <https://orcid.org/0000-0003-0108-6021>.

J.R. Kogan is associate dean, Student Success and Professional Development, and professor of medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; ORCID: <https://orcid.org/0000-0001-8426-9506>.

References

- Crosson FJ, Leu J, Roemer BM, Ross MN. Gaps in residency training should be addressed to better prepare doctors for a twenty-first century delivery system. *Health Aff.* 2011;30:2142–2148.
- Cooke M, Irby DM, O'Brien BC. *Educating Physicians: A Call for Reform of Medical School and Residency*. San Francisco, CA: Jossey-Bass; 2010.
- Lynn L, Hess BJ, Weng W, Lipner RS, Holmboe ES. Gaps in the quality of diabetes care in internal medicine residency clinics suggest the need for better ambulatory training. *Health Aff.* 2012;31:150–158.
- Wennberg JE. *Tracking Medicine: A Researcher's Quest to Understand Healthcare*. New York, NY: Oxford University Press; 2010.
- Wennberg JE, Fowler FJ Jr. A test of consumer contribution to small area variations in health care delivery. *J Maine Med Assoc.* 1977;68:275–279.
- The Dartmouth Atlas Project. The Dartmouth Atlas of Health Care. <https://www.dartmouthatlas.org>. Accessed March 1, 2022.
- Donabedian A. *An Introduction to Quality Assurance in Health Care*. New York, NY: Oxford University Press; 2003.
- Nelson EC, Batalden PB, Godfrey MM. *Quality by Design. A Clinical Microsystems Approach*. San Francisco, CA: Jossey Bass; 2007.
- Batalden P, Ovalle A, Foster T, Elwyn G. Science-informed practice: An essential contributor to health-care coproduction. *Int J Qual Health Care.* 2021;33:ii4–ii5.
- Agency for Healthcare Quality and Research. 2019 National Healthcare Quality and Disparities Report. <https://www.ahrq.gov/research/findings/nhqdr/nhqdr19/index.html>. Accessed March 1, 2022.
- Choosing Wisely. <https://www.choosingwisely.org>. Accessed March 2, 2022.
- Holmes-Rovner M, Montgomery JS, Rovner DR, et al. Informed decision making: Assessment of the quality of physician communication about prostate cancer diagnosis and treatment. *Med Decis Making.* 2015;35:999–1009.
- The Commonwealth Fund. Radley DC, Collins SR, Baumgartner JC. 2020 Scorecard on State Health System Performance. <https://www.commonwealthfund.org/publications/scorecard/2020/sep/2020-scorecard-state-health-system-performance>. Accessed March 2, 2022.
- Egede LE, Walker RJ. Structural racism, social risk factors, and Covid-19—A dangerous convergence for Black Americans. *N Engl J Med.* 2020;383:e77.
- Chen C, Petterson S, Phillips R, Bazemore A, Mullan F. Spending patterns in region of residency training and subsequent expenditures for care provided by practicing physicians for Medicare beneficiaries. *JAMA.* 2014;312:2385–2393.
- Wennberg JE, Freeman JL, Culp WJ. Are hospital services rationed in New Haven or over-utilised in Boston? *Lancet.* 1987;1:1185–1189.
- Centers for Medicare and Medicaid. Find and Compare Nursing Homes, Hospitals and Other Providers Near You. <https://www.medicare.gov/care-compare/?providerType=Hospital&redirect=true>. Accessed March 2, 2022.
- Asch DA, Nicholson S, Srinivas S, Herrin J, Epstein AJ. Evaluating obstetrical residency programs using patient outcomes. *JAMA.* 2009;302:1277–1283.
- Epstein AJ, Srinivas SK, Nicholson S, Herrin J, Asch DA. Association between physicians' experience after training and maternal obstetrical outcomes: Cohort study. *BMJ.* 2013;346:f1596.
- Bansal N, Simmons KD, Epstein AJ, Morris JB, Kelz RR. Using patient outcomes to evaluate general surgery residency program performance. *JAMA Surg.* 2016;151:111–119.
- Phillips RL Jr, Petterson SM, Bazemore AW, Wingrove P, Puffer JC. The effects of training institution practice costs, quality, and other characteristics on future practice. *Ann Fam Med.* 2017;15:140–148.
- The CLER Evaluation Committee and the CLER Program. CLER National Report of Findings 2019: Initial Visits to Sponsoring Institutions With 2 or Fewer Core Residency Programs. Chicago, IL: Accreditation Council for Graduate Medical Education; 2019. <https://www.acgme.org/globalassets/PDFs/CLER/ACGMECLERNationalReportFindings2019.pdf>. Accessed March 2, 2022.
- Holmboe E, Ginsburg S, Bernabeo E. The rotational approach to medical education: Time to confront our assumptions? *Med Educ.* 2011;45:69–80.
- Hunt D, Migdal M, Eaglen R, Barzansky B, Sabalis R. The unintended consequences of clarity: Reviewing the actions of the Liaison Committee on Medical Education before and after the reformatting of accreditation standards. *Acad Med.* 2012;87:560–566.
- Coalition for Physician Accountability Undergraduate Medical Education—Graduate Medical Education Review Committee (UGRC): Recommendations for Comprehensive Improvement of the UME–GME Transition. <https://physicianaccountability.org/wp-content/uploads/2021/08/UGRC-Coalition-Report-FINAL.pdf>. Accessed March 2, 2022.
- Bernabeo EC, Holtman MC, Ginsburg S, Rosenbaum JR, Holmboe ES. Lost in transition: The experience and impact of frequent changes in the inpatient learning environment. *Acad Med.* 2011;86:591–598.
- Denson JL, Jensen A, Saag HS, et al. Association between end-of-rotation resident transition in care and mortality among hospitalized patients. *JAMA.* 2016;316:2204–2213.
- Denson JL, McCarty M, Fang Y, Uppal A, Evans L. Increased mortality rates During resident handoff periods and the effect of ACGME duty hour regulations. *Am J Med.* 2015;128:994–1000.
- Denson JL, Knoeckel J, Kjerengtroen S, et al. Improving end-of-rotation transitions of care among ICU patients. *BMJ Qual Saf.* 2020;29:250–259.
- Pangaro L. “Forward feeding” about students' progress: More information will enable better policy. *Acad Med.* 2008;83:802–803.
- Pusic MV, Boutis K, Hatala R, Cook DA. Learning curves in health professions education. *Acad Med.* 2015;90:1034–1042.
- Sedlack RE, Coyle WJ; ACE Research Group. Assessment of competency in endoscopy: Establishing and validating generalizable competency benchmarks for colonoscopy. *Gastrointest Endosc.* 2016;83:516–23.e1.
- McGaghie WC. When I say ... mastery learning. *Med Educ.* 2015;49:558–559.
- Barsuk JH, Cohen ER, Wayne DB, McGaghie WC, Yudkowsky R. A comparison of approaches for mastery learning standard setting. *Acad Med.* 2018;93:1079–1084.
- American Association of Collegiate Registrars and Admissions Officers. *Family Educational Rights and Privacy Act*. <https://>

- www.aacrao.org/advocacy/compliance/ferpa. Accessed March 2, 2022.
- 36 Chassin MR, Loeb JM. The ongoing quality improvement journey: Next stop, high reliability. *Health Aff (Millwood)*. 2011;30:559–568.
 - 37 Smirnova A, Sebok-Syer SS, Chahine S, et al. Defining and adopting clinical performance measures in graduate medical education: Where are we now and where are we going? *Acad Med*. 2019;94:671–677.
 - 38 Holmboe ES, Iobst WF. Accreditation Council for Graduate Medical Education. Assessment Guidebook. <https://www.acgme.org/globalassets/PDFs/Milestones/Guidebooks/AssessmentGuidebook.pdf?ver=2020-11-18-155141-527&ver=2020-11-18-155141-527>. Accessed March 2, 2022.
 - 39 Andolsek K, Padmore J, Hauer KE, Ekpenyong A, Edgar L, Holmboe ES. Accreditation Council for Graduate Medical Education. Clinical Competency Committees: A Guidebook for Programs. 3rd ed. <https://www.acgme.org/globalassets/PDFs/Milestones/Guidebooks/AssessmentGuidebook.pdf?ver=2020-11-18-155141-527&ver=2020-11-18-155141-527>. Accessed March 2, 2022.
 - 40 Gingerich A, Kogan J, Yeates P, Govaerts M, Holmboe E. Seeing the ‘black box’ differently: Assessor cognition from three research perspectives. *Med Educ*. 2014;48:1055–1068.
 - 41 Gingerich A, Regehr G, Eva KW. Rater-based assessments as social judgments: Rethinking the etiology of rater errors. *Acad Med*. 2011;86(10 suppl):S1–S7.
 - 42 Kogan JR, Hess BJ, Conforti LN, Holmboe ES. What drives faculty ratings of residents’ clinical skills? The impact of faculty’s own clinical skills. *Acad Med*. 2010;85(10 suppl):S25–S28.
 - 43 Barsuk JH, Cohen ER, Nguyen D, et al. Attending physician adherence to a 29-component central venous catheter bundle checklist during simulated procedures. *Crit Care Med*. 2016;44:1871–1881.
 - 44 Kogan JR, Conforti LN, Iobst WF, Holmboe ES. Reconceptualizing variable rater assessments as both an educational and clinical care problem. *Acad Med*. 2014;89:721–727.
 - 45 Kennedy TJ, Lingard L, Baker GR, Kitchen L, Regehr G. Clinical oversight: Conceptualizing the relationship between supervision and safety. *J Gen Intern Med*. 2007;22:1080–1085.
 - 46 Gingerich A, Daniels V, Farrell L, Olsen SR, Kennedy T, Hatala R. Beyond hands-on and hands-off: Supervisory approaches and entrustment on the inpatient ward. *Med Educ*. 2018;52:1028–1040.
 - 47 ten Cate O, Hart D, Ankel F, et al; International Competency-Based Medical Education Collaborators. Entrustment decision making in clinical training. *Acad Med*. 2016;91:191–198.
 - 48 Sheu L, Kogan JR, Hauer KE. How supervisor experience influences trust, supervision, and trainee learning: A qualitative study. *Acad Med*. 2017;92:1320–1327.
 - 49 Apramian T, Cristancho S, Sener A, Lingard L. How do thresholds of principle and preference influence surgeon assessments of learner performance? *Ann Surg*. 2018;268:385–390.
 - 50 Schumacher DJ, Holmboe E, Carraccio C, et al. Resident-sensitive quality measures in the pediatric emergency department: Exploring relationships with supervisor entrustment and patient acuity and complexity. *Acad Med*. 2020;95:1256–1264.
 - 51 Gennis VM, Gennis MA. Supervision in the outpatient clinic: Effects on teaching and patient care. *J Gen Intern Med*. 1993;8:378–380.
 - 52 Omori DM, O’Malley PG, Kroenke K, Landry F. The impact of the bedside visit in the ambulatory clinic. Does it make a difference? *J Gen Intern Med*. 1997;12(suppl 1):96a.
 - 53 Kogan JR, Conforti LN, Bernabeo E, Iobst W, Holmboe E. How faculty members experience workplace-based assessment rater training: A qualitative study. *Med Educ*. 2015;49:692–708.
 - 54 Singh H, Giardina TD, Forjuoh SN, et al. Electronic health record-based surveillance of diagnostic errors in primary care. *BMJ Qual Saf*. 2012;21:93–100.
 - 55 Selby JV, Schmittiel JA, Lee J, et al. Meaningful variation in performance: What does variation in quality tell us about improving quality? *Med Care*. 2010;48:133–139.
 - 56 Boonyasai RT, Windish DM, Chakraborti C, Feldman LS, Rubin HR, Bass EB. Effectiveness of teaching quality improvement to clinicians: A systematic review. *JAMA*. 2007;298:1023–1037.
 - 57 Johnson B. And: Making a Difference by Leveraging Polarity, Paradox or Dilemma. Amherst, MA: HRD Press; 2020.
 - 58 Batalden M, Batalden P, Margolis P, et al. Coproduction of healthcare service. *BMJ Qual Saf*. 2016;25:509–517.
 - 59 Englander R, Holmboe E, Batalden P, et al. Coproducing health professions education: A prerequisite to coproducing health care services? *Acad Med*. 2020;95:1006–1013.
 - 60 Wong BM, Holmboe ES. Transforming the academic faculty perspective in graduate medical education to better align educational and clinical outcomes. *Acad Med*. 2016;91:473–479.

Reference cited only in Figure 2

- 61 Levinson W, Lesser CS, Epstein RM. Developing physician communication skills for patient-centered care. *Health Aff (Millwood)*. 2010;29:1310–1318.