Revisiting the Core Entrustable Professional Activities for Entering Residency
Jonathan M. Amiel, MD, Dorothy A. Andriole, MD, Diane M. Biskobing, MD, David R. Brown, MD, William B. Cutrer, MD, MEd, Matthew T. Emery, MD, George C. Mejicano, MD, MS, Michael S. Ryan, MD, MEHP, Jennifer L. Swails, MD, and Dianne P. Wagner, MD, for the Association of American Medical Colleges Core EPAs for Entering Residency Pilot Team

Abstract

The Core EPAs for Entering Residency Pilot project aimed to test the feasibility of implementing 13 entrustable professional activities (EPAs) at 10 U.S. medical schools and to gauge whether the use of the Core EPAs could improve graduates’ performance early in residency. In this manuscript, the authors (members of the pilot institutions and Association of American Medical Colleges staff supporting the project evaluation) describe the schools’ capacity to collect multimodal evidence about their students’ performance in each of the Core EPAs and the ability of faculty committees to use those data to make decisions regarding learners’ readiness for entrustment. In reviewing data for each of the Core EPAs, the authors reflected on how each activity performed as an EPA informed by how well it could be assessed and entrusted. For EPAs that did not perform well, the authors examined whether there are underlying practical and/or theoretical issues limiting its utility as a measure of student performance in medical school.

Theoretical advances aiming to align medical education programs’ intended outcomes, instruction, learning, assessment, and progression to focus on a learner’s emerging ability to function in a professional environment through competency-based medical education (CBME), programmatic assessment, and entrustable professional activities (EPAs) have created a fertile ground for implementation pilots across the continuum of medical education.

Implementations in undergraduate medical education (UME) in particular have many face limitations in that in most settings, there is little learners may do without direct supervision beyond reviewing a medical record, taking a history and doing an initial physical examination despite the fact that first-day interns are, functionally, expected to perform many activities without direct supervision. In 2014, the Association of American Medical Colleges (AAMC) published a set of 13 core EPAs for entering residency (Core EPAs) that a first postgraduate year resident (PGY-1) could be expected to perform with indirect supervision on the first day of residency. That same year, the AAMC administered a survey to program directors (PDs) nationally (see complete survey, Supplemental Digital Appendix 1 at http://links.lww.com/ACADMED/B86). As shown in Figure 1, there was wide variation in the percentage of PDs that were confident that most or all of their PGY-1s (among those PGY-1s who were Liaison Committee for Medical Education–accredited medical school graduates) could perform each Core EPA without direct supervision in their first week of residency. These percentages ranged from 24.2% for EPA 12, “Perform the general procedures of a physician” to 88.4% for EPA 1, “Gather a medical history and perform a physical examination.”

In this manuscript, we describe selected outcomes of the implementation of the Core EPAs across pilot institutions. In doing so, we report data in aggregate from our implementations and provide our perspective on future directions for the Core EPAs in UME.

The Pilot

Shortly following the publication of the Core EPAs, the AAMC launched a pilot of 10 medical schools to determine the feasibility of implementing the Core EPAs and examining what was required to define, test, and iteratively improve UME EPAs to facilitate their use as a measure of students’ readiness to progress to residency.

The primary goal of the pilot was to design and implement educational processes that use the Core EPAs framework to optimize safe and effective patient care by ensuring that each graduate is prepared for the core initial duties of an entering resident. It had 2 aims:

- Share lessons learned regarding the framework of the Core EPAs to optimize this approach and foster its propagation throughout UME programs and
- Facilitate the transition from UME to graduate medical education (GME) via valid assessments of student trustworthiness, knowledge, skills, and attitudes as reflected in the EPAs.

From its outset, the pilot recognized its heterogeneity. Schools in the pilot were old and new, small and large, rural and urban, spread across the United States, variably resourced, in different stages of readiness for implementation of new curricular elements, and had variable expertise in CBME. Each institution committed to implementing at least 4 Core EPAs, with all of the Core EPAs represented across the pilot in aggregate across pilot sites.

doi: 10.1097/ACM.0000000000004088
Copyright © 2021 by the Association of American Medical Colleges
Supplemental digital content for this article is available at http://links.lww.com/ACADMED/B86.
The schools committed to following common guiding CBME principles, shown in List 1. To inform a common approach to assessment, developmental schema marking growth toward entrustment for each EPA were created. Between 2015 and 2017, each institution implemented its own set of assessments aligned to the developmental schema and the Core EPAs as originally published.

Data Collections
The collection of aggregate data allowed us to consider our ability to gather evidence about learner performance and readiness for entrustment in each of the Core EPAs that were implemented. Of the 10 schools in the pilot, 7 schools fulfilled both of the following 2 criteria for inclusion in the aggregated data presented in this report: 1) the school implemented at least 4 Core EPAs for all students in the incoming class of 2015 or 2016 and 2) the school participated in all Core EPAs pilot multischool studies for which the school was eligible. Three schools that did not meet both inclusion criteria are excluded from the aggregated data presented in this report. Data for the other 7 schools and their learners are summarized in this report. We examined various data for these schools' learners, accounting for each schools' time frame for implementation and the specific EPAs implemented (see table summarizing the variety of assessments used by these 7 schools for each EPA, Supplemental Digital Appendix 2 at http://links.lww.com/ACADMED/B86) collected from participating schools' learners directly by the AAMC or by the participating schools (see a detailed description of each data collection, Supplemental Digital Appendix 3 at http://links.lww.com/ACADMED/B86).

Data collected directly by the AAMC included:
1. Self-reported frequency of observation/feedback in performing the activities of the Core EPAs in the clinical workplace during medical school, as reported by students on the AAMC Medical School Graduation Questionnaire (GQ) in 2019 and 2020 (activities were not identified as “Core EPAs” on the questionnaire). 9–11
2. Self-reported readiness to perform the activities of the Core EPAs under indirect supervision, as reported by class of 2019 graduates on the AAMC Early Postgraduate Year (PGY)-1 Questionnaire (activities were not identified as “Core EPAs” on the questionnaire).

Data collected by schools included:
1. Numbers of EPA-specific workplace-based assessments (WBAs) completed for students between 2017 and 2020, as reported by each school to coauthor (M.S.R.).
2. Outcomes of trained entrustment group (TEG) deliberations: Pilot schools' TEGs assessed their students' readiness for entrustment and rendered "mock" (i.e., for program evaluation/process improvement purposes only) summative entrustment decisions for their graduates, considering either the entire class or a randomly selected subset of the class.12
3. PD assessment of graduates' preparedness: Pilot schools sent questionnaires to PDs about the readiness of 2019 graduates to perform Core EPAs under indirect supervision at the start of training. Summary, de-identified data for each school were aggregated by the first author (J.M.A.).

Figure 1 The AAMC 13 Core Entrustable Professional Activities (EPAs) for entering residency and 2014 AAMC program director EPAs survey results. Percentage of program directors “confident for most or all” = percentage of program directors who responded “most” or “all” to the following item on the survey: “Considering only the PGY-1 residents in your program who are 2014 graduates of LCME-accredited U.S. medical schools, please indicate how many residents you are confident were prepared to do the following without direct supervision in the first week of residency:” (response choices: “No or few,” “some,” “most,” or “all”). Total N = 503 program directors. Abbreviations: AAMC, Association of American Medical Colleges; EPA, entrustable professional activity; PGY, postgraduate year; LCME, Liaison Committee on Medical Education.
List 1

Core EPAs Pilot Guiding Principles

- Employ a systematic approach to map educational opportunities and assessments for each EPA
- Explicitly measure the attributes of trustworthiness in addition to the specific knowledge, skills, and attitudes required for each EPA
- Create a longitudinal view of each learner’s performance via, at minimum, aggregated performance evidence, and consider the added value of longitudinal relationships and formal coaching structures in informing entrustment decisions
- Gather multimodal performance evidence from multiple assessors about each learner for each EPA
- Include global professional judgments about the entrustment of each learner in the body of evidence that supports summative entrustment decisions
- Ensure a process for formative feedback along the trajectory to entrustment to provide opportunities for both remediation and potential acceleration of responsibilities
- Create a process to render and maintain formal entrustment decisions by a trained group (entrustment committee) that reviews performance evidence for each learner
- Ensure that each learner is an active participant in the entrustment process—aware of expectations, engaged in gathering and reviewing performance evidence, and generating individualized learning plans to attain entrustment
- Align formal entrustment decisions regarding individual learners with nationally established performance expectations, as currently described in the Core EPAs Curriculum Developer’s Guide

Abbreviation: EPA, entrustable professional activity.

Findings

In reflecting on the performance of each EPA, the authors adapted standards from the Joint Committee on Standards for Educational Evaluation including propriety, feasibility, utility, and accuracy. Propriety standards focus on the regard for the welfare of those affected by the evaluation. Feasibility standards focus on the practicality, prudence, and resource use of the evaluation. Utility standards focus on an evaluation’s ability to serve the needs of its users. Accuracy standards focus on the validity and reliability of the information gathered.

Propriety

The pilot recognized that the Core EPAs had never been tested empirically among cohorts of entire classes of medical students and that their implementation at our institutions would be the most significant foray into CBME that our schools had undertaken to date. All 10 schools in the pilot took a “first do no harm” approach, ensuring that entrustment decisions would not play a role in progression and that information about student performance would be used strictly for formative purposes as part of the pilot. We also adopted a skeptical approach, challenging ourselves to see whether, in the absence of the infusion of external resources, we would be able to build in enough curriculum, assessment, coaching, and data management to facilitate a robust implementation of the Core EPAs at our institutions. In doing so, we aimed to minimize potential tensions identified in the implementation of the Core EPAs at our institutions so that our learners’ education would only be potentially enhanced by our implementation.

Feasibility

The feasibility of implementing each of the Core EPAs was a major focus of all 10 schools in the pilot. The predominant method intended for assessment of all EPAs was WBA requiring direct observation in clinical settings. After being observed performing a specific EPA, a supervisor rates the student’s need for assistance and/or their readiness for practice with less supervision using one of several available supervisory scales.

Measures of extent of data generated (as available, from among the 7 schools represented in the aggregate data collections) are as follows:

- The mean number of WBAs completed per student ranged across EPAs from < 1.0 to 7.9 (Table 1, column A). For EPA 1 (History and physical, H&P), EPA 2 (Differential diagnosis, DDx), and EPA 6 (Oral presentation), the mean was > 5.0; for all other EPAs, the mean ranged from < 1.0 to 3.5.

Proportions of GQ respondents who reported observation/feedback (although not necessarily in written, WBA format) > 5 times ranged across EPAs from 7% to 95%, including > 75% of respondents for 7 of the 13 EPAs, 50% to 75% of respondents for 3 EPAs, and < 50% of respondents for 3 EPAs (Table 1, column B).

Among students considered by TEGs, proportions of students with ≥ 4 WBAs ranged across EPAs from 0% to 86%, including > 75% of students for EPA 2 (DDx), 50% to 75% of students for EPA 7 (EBM) and EPA 9 (Interprofessional collaboration), and < 50% of students for the remaining 10 EPAs (Table 1, column C).

Proportions of students for whom TEGs made an entrustment decision (versus did not make an entrustment decision, generally due to data limitations of quality/quantity/consistency) ranged across EPAs from 42% to 94%. TEGs made entrustment decisions (i.e., one of “ready for entrustment under indirect supervision,” “progressing toward readiness for entrustment,” “evidence is against progress toward readiness”) for > 75% of students for 9 EPAs, 50%–75% of students for 3 EPAs, and < 50% of students for EPA 11 (Informed consent) (Table 1, column D).

Utility

Measures of readiness for entrustment under indirect supervision, the intended outcome of the implementation activities, are shown (as available, from among the 7 schools represented in the aggregate data collections), as follows:

Among all students considered by TEGs, proportions deemed ready for entrustment under indirect supervision ranged across EPAs from 0% to 78%, including > 75% of students for EPA 6 (Oral presentation), 50% to 75% of students for 6 EPAs, and < 50% of students for the remaining 6 EPAs (Table 2, column A).

Among all students considered by TEGs, proportions deemed ready for entrustment under indirect supervision with moderate–high confidence on the part of the TEGs ranged across EPAs from 0% to 78%, including > 75% of students for EPA 6 (Oral presentation); 50% to 75% of students for EPA 1 (H&P), EPA 2 (DDx), EPA 5 (Documentation), EPA 7 (EBM), and EPA 9 (Interprofessional collaboration); and < 50% of students
for EPA 3 (Recommend/interpret tests), EPA 4 (Orders), EPA 8 (Handovers), EPA 10 (Procedures), EPA 11 (Informed consent), EPA 12 (Procedures), and EPA 13 (Safety) (Table 2, column B).

Proportions of graduates who assessed themselves as having been ready to perform the EPA under indirect supervision at the start of residency ranged across EPAs from 32% to 100%, including > 75% of graduates for EPA 1 (H&P), EPA 2 (DDx), EPA 5 (Documentation), EPA 6 (Oral presentation), EPA 7 (EBM), EPA 9 (Interprofessional collaboration), EPA 11 (Informed consent), and EPA 13 (Safety); 50% to 75% of graduates for EPA 3 (Recommend/interpret tests), EPA 4 (Orders), EPA 8 (Handovers), and EPA 10 (Urgent care); and < 50% of graduates for EPA 12 (Procedures) (Table 2, column C). These data were collected from graduates after completion of at least 3 months of PGY-1 training; responses may reflect the impact of residency orientation activities or bootcamps after medical school graduation, rather than any impact of their medical school activities alone.

Proportions of graduates that PDs rated as prepared to perform the EPA (see exact item wording, Supplemental Digital Appendix 3 at http://links.lww.com/ACADMED/B86) ranged across EPAs from 69% to 93%, including > 75% of graduates for all EPAs except EPA 12 (Procedures) (PDs rated 69% of graduates as prepared to perform this EPA) (Table 2, column D). PDs provided these data after these graduates had completed at least 3 months of PGY-1 training; responses may reflect the impact of residency orientations.

### Table 1
Multisource Data: Data Generated by EPA*

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EPA 1</strong>: Gather a history and perform a physical examination</td>
<td>7.4</td>
<td>78 (946/1,217)</td>
<td>47 (342/732)</td>
</tr>
<tr>
<td><strong>EPA 2</strong>: Prioritize a differential diagnosis following a clinical encounter</td>
<td>5.3</td>
<td>91 (686/756)</td>
<td>86 (255/296)</td>
</tr>
<tr>
<td><strong>EPA 3</strong>: Recommend and interpret common diagnostic and screening tests</td>
<td>3.6</td>
<td>92 (530/579)</td>
<td>46 (95/207)</td>
</tr>
<tr>
<td><strong>EPA 4</strong>: Enter and discuss orders and prescriptions</td>
<td>1.3</td>
<td>71 (334/469)</td>
<td>7 (15/207)</td>
</tr>
<tr>
<td><strong>EPA 5</strong>: Document a clinical encounter in the patient record</td>
<td>2.1</td>
<td>89 (648/730)</td>
<td>11 (50/452)</td>
</tr>
<tr>
<td><strong>EPA 6</strong>: Provide an oral presentation of a clinical encounter</td>
<td>7.9</td>
<td>95 (896/943)</td>
<td>48 (333/687)</td>
</tr>
<tr>
<td><strong>EPA 7</strong>: Form clinical questions and retrieve evidence to advance patient care</td>
<td>1.9</td>
<td>79 (594/752)</td>
<td>55 (228/416)</td>
</tr>
<tr>
<td><strong>EPA 8</strong>: Give or receive a patient handover to transition care responsibility</td>
<td>1.0</td>
<td>62 (460/738)</td>
<td>5 (14/277)</td>
</tr>
<tr>
<td><strong>EPA 9</strong>: Collaborate as a member of an interprofessional team</td>
<td>2.6</td>
<td>85 (728/857)</td>
<td>53 (221/416)</td>
</tr>
<tr>
<td><strong>EPA 10</strong>: Recognize a patient requiring urgent/emergent care and initiate evaluation/management</td>
<td>0.9</td>
<td>64 (274/427)</td>
<td>4 (8/182)</td>
</tr>
<tr>
<td><strong>EPA 11</strong>: Obtain informed consent for tests and/or procedures</td>
<td>0.8</td>
<td>36 (243/680)</td>
<td>&lt; 1 (1/302)</td>
</tr>
<tr>
<td><strong>EPA 12</strong>: Perform general procedures of a physician</td>
<td>0.2</td>
<td>19 (169/999)</td>
<td>4 (20/457)</td>
</tr>
<tr>
<td><strong>EPA 13</strong>: Identify system failures and contribute to a culture of safety and improvement</td>
<td>0.6</td>
<td>7 (66/991)</td>
<td>0 (0/252)</td>
</tr>
</tbody>
</table>

Abbreviations: AY, academic year; EPA, entrustable professional activity; WBA, workplace-based assessment; AAMC, Association of American Medical Colleges; GQ, Graduation Questionnaire; PGY, postgraduate year; TEG, trained entrustment group.

*Aggregate data in each column include data for those schools (up to 7 total) that had implemented the EPA for their students in the stated academic year(s). As schools varied in the EPAs that were implemented (from 4 to 13/school), denominators vary across EPAs within a given column of aggregated data. For Column A, table shading indicates: light gray—fewer than 2 WBA/student; medium gray—2–5 WBA/student; dark gray—5 or more WBA/student. For columns B–D, table shading indicates: light gray—< 50%; medium gray—50%–75%; dark gray—> 75%.

bn = 57,153 assessments for 1,764 learners (including those collected in the M3 and M4 years).

cWBAs data missing for one individual.

Data shown for procedure of “intravenous line placement.”

EPA 9 (Interprofessional collaboration), EPA 11 (Informed consent), and EPA 13 (Safety); 50% to 75% of graduates for EPA 3 (Recommend/interpret tests), EPA 4 (Orders), EPA 8 (Handovers), and EPA 10 (Urgent care); and < 50% of graduates for EPA 12 (Procedures) (Table 2, column C). These data were collected from graduates after completion of at least 3 months of PGY-1 training; responses may reflect the impact of residency orientation activities or bootcamps after medical school graduation, rather than any impact of their medical school activities alone. Proportions of graduates that PDs rated as prepared to perform the EPA (see exact item wording, Supplemental Digital Appendix 3 at http://links.lww.com/ACADMED/B86) ranged across EPAs from 69% to 93%, including > 75% of graduates for all EPAs except EPA 12 (Procedures) (PDs rated 69% of graduates as prepared to perform this EPA) (Table 2, column D). PDs provided these data after these graduates had completed at least 3 months of PGY-1 training; responses may reflect the impact of residency orientations.
Accuracy

Two major challenges were presented in our implementation of assessments to measure learner performance of the Core EPAs. First, before the initiation of the Core EPAs pilot, many institutions had already incorporated instruments to measure performance in one or more aspects of some Core EPAs. In these situations, developing a consensus around any one instrument became challenging. As a consequence, members of the pilot sought existing tools, particularly those that may be useful for direct observation in the workplace, to measure performance of the Core EPAs. This brought about the second challenge, namely the lack of validity and reliability data for instruments used for this purpose in this population.

Toward the middle of the pilot, following the recommendations of our supervisory scale task force, most institutions incorporated one or both of 2 scales to measure performance across Core EPAs in the workplace. 8 The modified Ottawa scale measures learner performance in a retrospective fashion based on the

Table 2

Multisource Data: Readiness for Entrustment by EPA*

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of schools in data collection by academic year (graduating class)</td>
<td>Among all students considered by the TEG, % (no.) of students deemed ready for entrustment in EPA</td>
<td>Among all students considered by the TEG, % (no.) of students deemed ready for entrustment in EPA with moderate–high confidence by TEG</td>
<td>Early PGY-1 questionnaire: % (no.) of graduates ready for indirect supervision at start of PGY-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPA 1. Gather a history and perform a physical examination</td>
<td>71 (518/732)</td>
<td>70 (513/732)</td>
<td>100 (98/98)</td>
</tr>
<tr>
<td>EPA 2. Prioritize a differential diagnosis following a clinical encounter</td>
<td>68 (200/296)</td>
<td>67 (198/296)</td>
<td>94 (45/48)</td>
</tr>
<tr>
<td>EPA 3. Recommend and interpret common diagnostic and screening tests</td>
<td>41 (84/207)</td>
<td>29 (60/207)</td>
<td>69 (20/29)</td>
</tr>
<tr>
<td>EPA 4. Enter and discuss orders and prescriptions</td>
<td>7 (15/207)</td>
<td>5 (10/207)</td>
<td>66 (19/29)</td>
</tr>
<tr>
<td>EPA 5. Document a clinical encounter in the patient record</td>
<td>64 (288/453)</td>
<td>53 (242/453)</td>
<td>95 (56/59)</td>
</tr>
<tr>
<td>EPA 6. Provide an oral presentation of a clinical encounter</td>
<td>78 (539/687)</td>
<td>78 (535/687)</td>
<td>94 (75/80)</td>
</tr>
<tr>
<td>EPA 7. Form clinical questions and retrieve evidence to advance patient care</td>
<td>74 (308/416)</td>
<td>51 (214/416)</td>
<td>98 (47/48)</td>
</tr>
<tr>
<td>EPA 8. Give or receive a patient handover to transition care responsibility</td>
<td>8 (21/277)</td>
<td>6 (17/277)</td>
<td>71 (35/49)</td>
</tr>
<tr>
<td>EPA 9. Collaborate as a member of an interprofessional team</td>
<td>57 (236/416)</td>
<td>50 (210/416)</td>
<td>90 (60/67)</td>
</tr>
<tr>
<td>EPA 10. Recognize a patient requiring urgent/emergent care and initiate evaluation/management</td>
<td>3 (5/182)</td>
<td>2 (3/182)</td>
<td>54 (15/28)</td>
</tr>
<tr>
<td>EPA 11. Obtain informed consent for tests and/or procedures</td>
<td>&lt; 1 (1/302)</td>
<td>0 (0/302)</td>
<td>78 (39/50)</td>
</tr>
<tr>
<td>EPA 12. Perform general procedures of a physician</td>
<td>54* (245/457)</td>
<td>0 (0/457)</td>
<td>32* (25/79)</td>
</tr>
<tr>
<td>EPA 13. Identify system failures and contribute to a culture of safety and improvement</td>
<td>0 (0/252)</td>
<td>0% (0/252)</td>
<td>76 (51/67)</td>
</tr>
</tbody>
</table>

Abbreviations: AY, academic year; EPA, entrustable professional activity; TEG, trained entrustment group; PGY, postgraduate year.

*Aggregate data in each column include data for those schools (up to 7 total) that had implemented the EPA for their graduates in the stated academic year(s). As schools varied in the EPAs that were implemented (from 4–13/school), denominators vary across EPAs within a given column of aggregated data. Table shading indicates: light gray—< 50%; medium gray—50%–75%; dark gray—> 75%.

*Of note, the majority of entrustments in EPA 12 were made based on data from simulation rather than authentic practice in the workplace and were noted to have low confidence due to the lack of WBAs (see below).

*Data shown for procedure of “intravenous line placement.”

activities or bootcamps in which these graduates participated after medical school graduation, rather than the impact of their medical school activities alone as well a number of months of PGY-1 training.
amount of intervention a supervisor needed to provide for a given EPA.\textsuperscript{15} The modified Chen scale measures supervision in a prospective fashion based on the supervisor's impression of how much supervision may be required in the future.\textsuperscript{2}

Collectively, pilot team members across all 10 participating pilot schools conducted several studies to assess the validity and reliability of the 2 respective scales. At Vanderbilt, a mixed methods study of both scales in fourth-year acting internships within pediatrics and internal medicine found that the scores obtained from these 2 scales were significantly different and not related.\textsuperscript{16} At Virginia Commonwealth University (VCU), generalizability studies found the variance attributed to learner (a desirable feature) was poor across the Core EPAs assessed using both the Ottawa and Chen scales.\textsuperscript{17} At Yale, where a smaller cohort of trained assessors used the scales, generalizability studies using the Ottawa scale in the internal medicine clerkship found that variance attributed to the learner was substantially greater than that found in the VCU-based study.\textsuperscript{18} In addition to these studies, the Core EPAs pilot studied the performance of both scales in a simulated setting using a series of video and paper-based vignettes. In this context, both scales performed relatively well in terms of inter-rater reliability and generalizability.\textsuperscript{17,18}

At the pilot schools that convened TEGs, these TEGs were convened for program evaluation/process improvement purposes only; entrustment decisions made by these TEGs have not yet been examined for evidence of concurrent or predictive validity.

Summary of findings

In summary, data regarding the accuracy of assessment in the Core EPAs were largely restricted to WBAs and were limited to several single-institutional studies and one simulation–based study. Overall, the data were mixed in terms of the validity and reliability of the scales proposed to measure performance of the Core EPAs. We suspect that some differences between institutions relate to challenges associated with faculty development and the volume of raters involved in training across clerkships and institutions. The validity and reliability of data obtained from assessments outside of WBAs that incorporated the Ottawa or Chen scales have not yet been established, nor have evidence of concurrent or predictive validity of TEG entrustment decisions. Finally, data available from graduates themselves during residency, and from their PDs, were very limited in scope.

In total, examining the standards of propriety, feasibility, utility, and accuracy applied by the authors, we found that the pilot met the propriety standard uniformly, but the marked heterogeneity in the data suggested that the framework does not yet meet standards for feasibility, utility, or accuracy. This was not surprising, since in our pilot, and in most other CBME implementations, the training changes did not occur with commensurate, necessary, systems change that would enable the work of the learners to substantiate their progression.

Revisiting the List of the 13 Core EPAs

In reflecting on our experience implementing the Core EPAs in our institutions over the past 6 years, we have found and described variability in the feasibility, utility, and accuracy of different EPAs. This variability is instructive and leads us to propose that there are 3 clusters of EPAs within the published 13 and a potential fourth cluster that includes additional EPAs not originally described in the AAMC's list of 13 Core EPAs.

The first cluster includes those EPAs that align well with existing curricula, meaning that there exist ample opportunities for learners to practice them with direct observation and feedback. We call these EPAs the “core of the Core,” and they include EPA 1 (H&P), EPA 2 (DXs), EPA 5 (Documentation), EPA 6 (Oral presentation), EPA 7 (EBM), and EPA 9 (Interprofessional collaboration). With these activities well represented in the UME experience, implementing them as EPAs is a powerful means to make them core to students’ educational experience and, in turn, generating sufficient data to inform entrustment. This approach can also potentially facilitate efforts with health systems to gradually withdraw supervision when a learner meets criteria for entrustment.\textsuperscript{19,20} For some of the EPAs in this cluster, entrustment may be feasible sooner than graduation, and may allow schools to permit students to take on certain patient care responsibilities under indirect supervision that in current systems are not permitted explicitly.

The second cluster are those EPAs that may be represented most prominently in the senior UME curricula, where learners may have opportunities to practice them, but perhaps in limited volume, with supervision that may be inconsistent or not sufficiently intentional to collect evidence robust enough to substantiate entrustment decisions. That we are challenged to entrust our graduates to perform EPA 3 (Tests), EPA 4 (Orders), or EPA 8 (Handovers) should raise concerns for us all, since these are activities that they will likely be expected to perform without direct supervision on their first day of residency. The experiences with these EPAs present a clear call for quality improvement in both curricula and assessment.

The third cluster includes EPAs that in most of our participating schools’ UME curricula appear to be absent or are meaningfully, and unfortunately, underdeveloped. We have called these EPAs the “aspirational EPAs” for our pilot, and they include EPA 10 (Urgent care), EPA 11 (Informed consent), EPA 12 (Procedures), and EPA 13 (Safety). Though critically important to the function of physicians, in our view, our data suggest that these EPAs were not authentic to the role of medical students in our health systems as they presently exist. Based on our pilot experiences and reports of others in the literature, EPA 12 (Procedures) in particular may be better aligned with the training of most resident physicians and of selected medical students in their final year based on the specialty they will enter for PGY-1 training, especially given that there may be specialty-specific differences in the extent to which PGY-1 residents are expected to perform various procedures without direct supervision at the start of training.\textsuperscript{20–23} According to results of a national survey of internal medicine PDs, the majority of PDs reported that new interns must/should be able to perform all Core EPAs except EPA 12 (Procedures), under indirect supervision at the start of residency.\textsuperscript{23} In the Education in Pediatrics Across the Continuum experience with authentic implementation of the Core EPAs (i.e., learners were not advanced...
to GME until they were determined to be ready for entrustment in the Core EPAs). EPA 12 (Procedures) was removed from this list as there were not sufficient opportunities for these students to practice these procedures. Finally, our pilot experiences suggest that the remaining EPAs in this third cluster—EPA 10 (Urgent care), EPA 11 (Informed consent), and EPA 13 (Safety)—require material investment in making them authentic activities of medical students and expanding the quantity and quality of assessment data that are collected to make them feasible for entrustment in UME.

We also acknowledge that there is likely an emerging fourth cluster of Core EPAs: those that were not included in the originally published list but are nevertheless core to the medical student role and may be considered for inclusion in UME CBME curricula. These could include activities related to educating patients about their illnesses, counseling them on their options, providing motivational interviewing, or providing telemedicine services.

There are important limitations in the aggregate data (aggregated across schools and across years) we provided in this report. Response rates for several surveys (e.g., PD surveys and Early PGY-1 survey) were low, and PD surveys differed across schools regarding both item wording and timing. We also note that both GQ data and Early PGY-1 data are self-report. Furthermore, as has been previously described, pilot schools differed in multiple aspects of the pilot implementation itself 6,10,22 that may contribute substantively to successful implementation, such as institutional investments in curriculum development, technology infrastructure to track learner performance, opportunities for deliberate and authentic practice, and workplace-based supervision. Nonetheless, these aggregate data that highlight numerous EPA-specific challenges in implementation for the schools in the AAMC Core EPAs pilot project may be of interest to the many U.S. medical schools that have incorporated the AAMC Core EPAs in their programmatic objectives.

Conclusions
The work of the Core EPAs pilot found a different subset of feasible EPAs for entering residents than the full set of those originally proposed or the subset supported by EPA developers. To date, the EPAs that were most successfully implemented at our pilot schools were those EPAs that were readily available for practice and assessment in UME curricula. Success, in this context, is an EPA that was authentically present in the defined role of medical students and thus leaned itself to observation, feedback, and eventual review of multisource data collected longitudinally for entrustment. Generally, EPAs that were less successfully implemented at our pilot schools represent gaps in the scope of practice of medical students at our institutions that must be remedied because they are truly core activities of first-day residents. Other EPAs that were less successfully implemented, such as EPA 12: Procedures, may be more appropriate for GME, or selectively in UME, based on specialty plans, in our view.

We suspect that those EPAs that were hampered by limitations in feasibility and/or utility in our pilot were those in which health systems have not created appropriate and necessary opportunities to entrust students. This represents a different opportunity or challenge from what we had anticipated when setting out in this work. Namely, to truly implement the Core EPA framework, the clinical environment—and its associated practical, legal, fiscal, and regulatory environments—must change to permit and support undergraduate educational experiences in which students carry greater responsibility for their patients than the modest scope of practice they currently have in most American medical schools. Curricula based on Core EPAs could then authentically inform progression.

We may have a “chicken and egg” problem here in that health systems, payors, and regulators may need more evidence, perhaps in the form of gateways toward increasing independence, 36 that medical students can safely manage their responsibilities and so a conversation between medical schools and their clinical affiliates needs to occur in which they, together, answer the question, “what evidence is required to substantiate true entrustment for patient care in this domain.” Once there is a shared mental model of the activities for which students can be entrusted, the design of the EPA-oriented curriculum may be tailored and adequate opportunities for observation and assessment may be planned. Finally, as the AAMC Core EPAs are intended to describe the activities that a PGY-1 resident is expected to perform with indirect supervision on the first day of residency, close collaboration with GME PDs across all specialties is essential in considering revisions to the initial list of AAMC Core EPAs and in further defining the role of the Core EPAs in optimally preparing medical school graduates for the responsibilities they assume at the start of residency training.

Acknowledgments: The authors wish to thank Douglas Grbic, PhD, lead research analyst; Amy Jayas, MPH, research analyst; Marie Caulfield, PhD, manager, applicant, student, and resident data; and David Matthews, PhD, senior research and data analyst, all at the Association of American Medical Colleges for provision of data.

Funding/Support: This work was supported by the Association of American Medical Colleges and by the medical schools participating in the Core Entrustable Professional Activities for Entering Residency pilot.

Other disclosures: Columbia University receives funding from the Association of American Medical Colleges in support of the work of J.M. Amiel related to administration of the pilot. Virginia Commonwealth University receives funding from the Association of American Medical Colleges in support of the work of M.S. Ryan as a consultant on graduate medical education.

Ethical approval: Reported as not applicable.

J.M. Amiel is associate professor of psychiatry, senior associate dean for curricular affairs, and interim co-vice dean for education, Columbia University Vagelos College of Physicians & Surgeons, New York, New York; ORCID: https://orcid.org/0000-0003-1538-9779.

D.A. Andriole is senior director, Medical Education Research, Association of American Medical Colleges, Washington, DC; ORCID: https://orcid.org/0000-0001-8902-1227.

D.M. Biskobing is professor of medicine and associate dean for pre-clinical medical education, Virginia Commonwealth University School of Medicine, Richmond, Virginia.

D.R. Brown is associate professor, chief, Family and Community Medicine, and interim chair, Humanities, Health, and Society, Florida International University Herbert Wertheim College of Medicine, Miami, Florida; ORCID: https://orcid.org/0000-0002-5361-6664.

W.B. Cutrer is associate dean for undergraduate medical education and associate professor of pediatrics, Critical Care Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee; ORCID: https://orcid.org/0000-0003-1538-9779.

M.T. Emery is associate professor of emergency medicine and medical director for simulation, Michigan State University College of Human Medicine, East Lansing, Michigan.
References


25. Engle B, Brogan-Hartlieb K, Obeso V, Pedoussaut M, Esposito K, Brown D. From the Classroom to Entrustment. The Development of Motivational Interviewing as an Entrustable Professional Activity. Presented the annual meeting of the Association for Medical Education in Europe; August 2020. Virtual.


Copyright © by the Association of American Medical Colleges. Unauthorized reproduction of this article is prohibited.