Title

Relationship between epa level of supervision with their associated subcompetency milestone levels in pediatric fellow assessment.

Permalink

https://escholarship.org/uc/item/4f50z83s

Journal

BMC Medical Education, 23(1)

Authors

Carraccio, Carol
Herman, Bruce
Weiss, Pnina
et al.

Publication Date

2023-10-03

DOI

10.1186/s12909-023-04689-0

Peer reviewed

RESEARCH

Open Access

Relationship between epa level of supervision with their associated subcompetency milestone levels in pediatric fellow assessment

Richard B. Mink[1*], Carol L. Carraccio[2], Bruce E. Herman[3], Pnina Weiss[4], David A. Turner[2], Diane E. J. Stafford[5], Kathleen A. McGann[6], Jennifer Kesselheim[7], Deborah C. Hsu[8], Pamela C. High[9,10], Jill J. Fussell[11], Megan L. Curran[12], Patricia R. Chess[13], Cary Sauer[14], Sarah Pitts[15], Angela L. Myers[16], John D. Mahan[17], Christiane E. L. Dammann[18], Tandy Aye[5], Alan Schwartz[19] and for the Subspecialty Pediatrics Investigator Network

Abstract

Background Entrustable Professional Activities (EPA) and competencies represent components of a competency-based education framework. EPAs are assessed based on the level of supervision (LOS) necessary to perform the activity safely and effectively. The broad competencies, broken down into narrower subcompetencies, are assessed using milestones, observable behaviors of one's abilities along a developmental spectrum. Integration of the two methods, accomplished by mapping the most relevant subcompetencies to each EPA, may provide a cross check between the two forms of assessment and uncover those subcompetencies that have the greatest influence on the EPA assessment.

Objectives We hypothesized that 1) there would be a strong correlation between EPA LOS ratings with the milestone levels for the subcompetencies mapped to the EPA; 2) some subcompetencies would be more critical in determining entrustment decisions than others, and 3) the correlation would be weaker if the analysis included only milestones reported to the Accreditation Council for Graduate Medical Education (ACGME).

Methods In fall 2014 and spring 2015, the Subspecialty Pediatrics Investigator Network asked Clinical Competency Committees to assign milestone levels to each trainee enrolled in a pediatric fellowship for all subcompetencies mapped to 6 Common Pediatric Subspecialty EPAs as well as provide a rating for each EPA based upon a 5-point LOS scale.

Results One-thousand forty fellows were assessed in fall and 1048 in spring, representing about 27% of all fellows. For each EPA and in both periods, the average milestone level was highly correlated with LOS (rho range 0.59–0.74; $p < 0.001$). Correlations were similar when using a weighted versus unweighted milestone score or using only the ACGME reported milestones ($p > 0.05$).

---

*Correspondence:
Richard B. Mink
rmink@ucla.edu
Full list of author information is available at the end of the article

Mink *et al. BMC Medical Education*   (2023) 23:720

Page 2 of 14

Conclusions We found a strong relationship between milestone level and EPA LOS rating but no difference if the subcompetencies were weighted, or if only milestones reported to the ACGME were used. Our results suggest that representative behaviors needed to effectively perform the EPA, such as key subcompetencies and milestones, allow for future language adaptations while still supporting the current model of assessment. In addition, these data provide additional validity evidence for using these complementary tools in building a program of assessment.

Keywords EPAs, Milestones, Assessment, Correlation, Fellow, Equations

Background

Early in the transition to a competency-based model for trainee education and assessment, identifying the Accreditation Council for Graduate Medical Education (ACGME) core competencies in the United States (US) and the CanMeds roles in Canada were critical first steps [1, 2]. Each of the core competencies was further refined into specific "subcompetencies" in the US and each Can-Meds role elaborated and defined. An important next step was the creation of milestones specific to the sub-competency or the role [3, 4]. The milestones represent defined, observable abilities of an individual's skills along a developmental continuum [5]. In the US, each specialty was tasked with creating both the subcompetencies and milestones for the ACGME competencies [3]. Pediatrics created milestones for 48 subcompetencies, of which only 21 were reported to the Accreditation Council for Graduate Medical Education (ACGME) biannually for all trainees [6]. Milestone ratings ranged from one to four or one to five, but trainees were not necessarily expected to achieve the highest levels at the time of graduation.

The subsequent creation of Entrustable Professional Activities (EPAs) by ten Cate and Scheele [7] complements the milestones by providing a meaningful clinical context for the subcompetencies. EPAs are observable activities of a profession that an individual should be able to execute without supervision when in practice [8–11]. As opposed to subcompetencies, EPA assessments are based upon the amount of supervision a trainee needs to perform the activity safely and effectively, ranging from direct to indirect to none [12, 13]. Basing EPA judgements on needed levels of trainee supervision aligns what faculty do in real time with what they are asked to do as part of the assessment process, thus adding to their validity evidence [14].

To link EPA and milestone assessments, medical education leaders then mapped the subcompetencies thought to be critical in executing the professional activities of each EPA [7, 15, 16]. An example of the mapping of the Leadteam EPA (Table 1) is illustrated in Fig. 1. For this EPA, 8 subcompetencies were judged to be important in making the entrustment decision. Milestones for 5 of the 8 subcompetencies are required to be reported to the ACGME in the fall and spring each year. While mapping was accomplished by experts through an iterative process, data supporting the mapping are lacking and it is unknown if any specific subcompetency is more important than the others in making the entrustment decision. Similarly, it is unclear whether using all the mapped subcompetencies, or only those required to be reported to the ACGME, are critical in formulating the entrustment decision. This information would be helpful to know for future studies, particularly if the milestone levels could be obtained directly from the ACGME.

Milestones and EPA level of supervision (LOS) both represent elements of a system of trainee assessment [17]. While all ACGME accredited programs must report milestones, many specialties are now promoting the use of EPA LOS for assessment. The American Board of Pediatrics announced that it will begin using EPA LOS ratings to determine eligibility to sit for its certification exams [18]. In July 2023, the American Board of Surgery began using EPAs as the foundation for competency-based surgical training [19]. Both Emergency Medicine and Family Medicine, along with other disciplines, are also currently exploring the use of EPA LOS in their training programs [18]. Since EPAs provide a holistic view ("wide lens" in Fig. 1) of the execution of an activity, and milestones provide a more granular assessment ("narrow lens") of specific behaviors needed to perform them, there should be a strong relationship between these two approaches to assessment [20]. This relationship requires further exploration as the finding of a strong association between the two would provide validity evidence for both types of assessments.

Using this logic, our hypotheses were: 1) there is a strong correlation between EPA LOS rating with the average score of the mapped subcompetency milestone levels needed to perform the EPA; 2) some subcompetencies would be more critical than others such that weighted scores will have a stronger correlation; and 3) if only those milestones required for reporting to the ACGME were included in the analysis, the correlation between EPA LOS rating and average milestone level would be weaker.

Mink *et al. BMC Medical Education*   (2023) 23:720

Page 3 of 14

Table 1 The six Common Pediatric Subspecialty EPAs evaluated in this study with their abbreviations and scales for EPA level of supervision

Common Pediatric Subspecialty EPA	Abbreviation
Apply public health principles and quality improvement methods to improve population health	**QI**
Level 1: Trusted to observe only	
Level 2: Trusted to contribute with direct supervision and coaching as a member of a collaborative effort to improve care at the institutional level	
Level 3: Trusted to contribute without direct coaching as a member of a collaborative effort to improve care at the institutional level	
Level 4: Trusted to lead collaborative efforts to improve care for populations and systems at the institutional level	
Level 5: Trusted to lead collaborative efforts to improve care at the level of populations and systems at the regional and/or national level	
Provide consultation to other healthcare providers caring for children and adolescents and refer patients requiring further consultation to other subspecialty providers if necessary	**Consultation**
Level 1:Trusted to observe only	
Level 2:Trusted to execute with direct supervision and coaching	
Level 3:Trusted to execute with indirect supervision and discussion of information conveyed for most simple and some complex cases	
Level 4:Trusted to execute with indirect supervision and may require discussion of information conveyed but only for selected complex cases	
Level 5:Trusted to execute independently without supervision	
Contribute to the fiscally sound, equitable and collaborative management of a healthcare workplace	**Management**
Level 1:Trusted to observe only	
Level 2:Trusted to perform with direct supervision and coaching with supervisor verifying work product for accuracy	
Level 3:Trusted to perform with supervisor serving as a consultant for all tasks	
Level 4:Trusted to perform with supervisor serving as a consultant but only for complex tasks	
Level 5:Trusted to perform without supervision	
Facilitate handovers to another healthcare provider either within or across settings	**Handover**
Level 1:Trusted to observe only	
Level 2:Trusted to execute with direct supervision and coaching	
Level 3:Trusted to execute with indirect supervision with verification of information after the handover for most simple and some complex cases	
Level 4:Trusted to execute with indirect supervision with verification of information after the handover for selected complex cases	
Level 5:Trusted to execute without supervision	
Lead an interprofessional health care team	**Leadteam**
Level 1:Trusted to participate only	
Level 2:Trusted to lead with direct supervision and coaching	
Level 3:Trusted to lead with supervisor occasionally present to provide advice	
Level 4:Trusted to lead without supervisor present but requires coaching to improve member and team performance	
Level 5:Trusted to lead without supervision to improve member and team performance	
Lead within the subspecialty profession	**Leadprof**
Level 1:Trusted to observe only	
Level 2:Trusted to contribute to advocacy and public education activities for the subspecialty profession with direct supervision and coaching at the institutional level	
Level 3:Trusted to contribute to advocacy and public education activities for the subspecialty profession with indirect supervision at the institutional level	
Level 4:Trusted to mentor others and lead advocacy and public education activities for the subspecialty profession at the institutional level	
Level 5:Trusted to lead advocacy and public education activities for the subspecialty profession at the regional and/or national level	

Methods

We performed the study using the Subspecialty Pediatrics Investigator Network (SPIN), a medical education research network that includes representatives from each of the 14 pediatric subspecialties with primary American Board of Pediatrics certification as well as the Council of Pediatric Subspecialties, Association of Pediatric Program Directors Fellowship Executive Committee, and Association of Pediatric Program Directors Longitudinal Educational Assessment Research Network (APPD LEARN) [21]. The goal was to recruit at least 20% of fellowships from each subspecialty. We obtained IRB approval from each participating institution with the University of Utah serving as the lead.
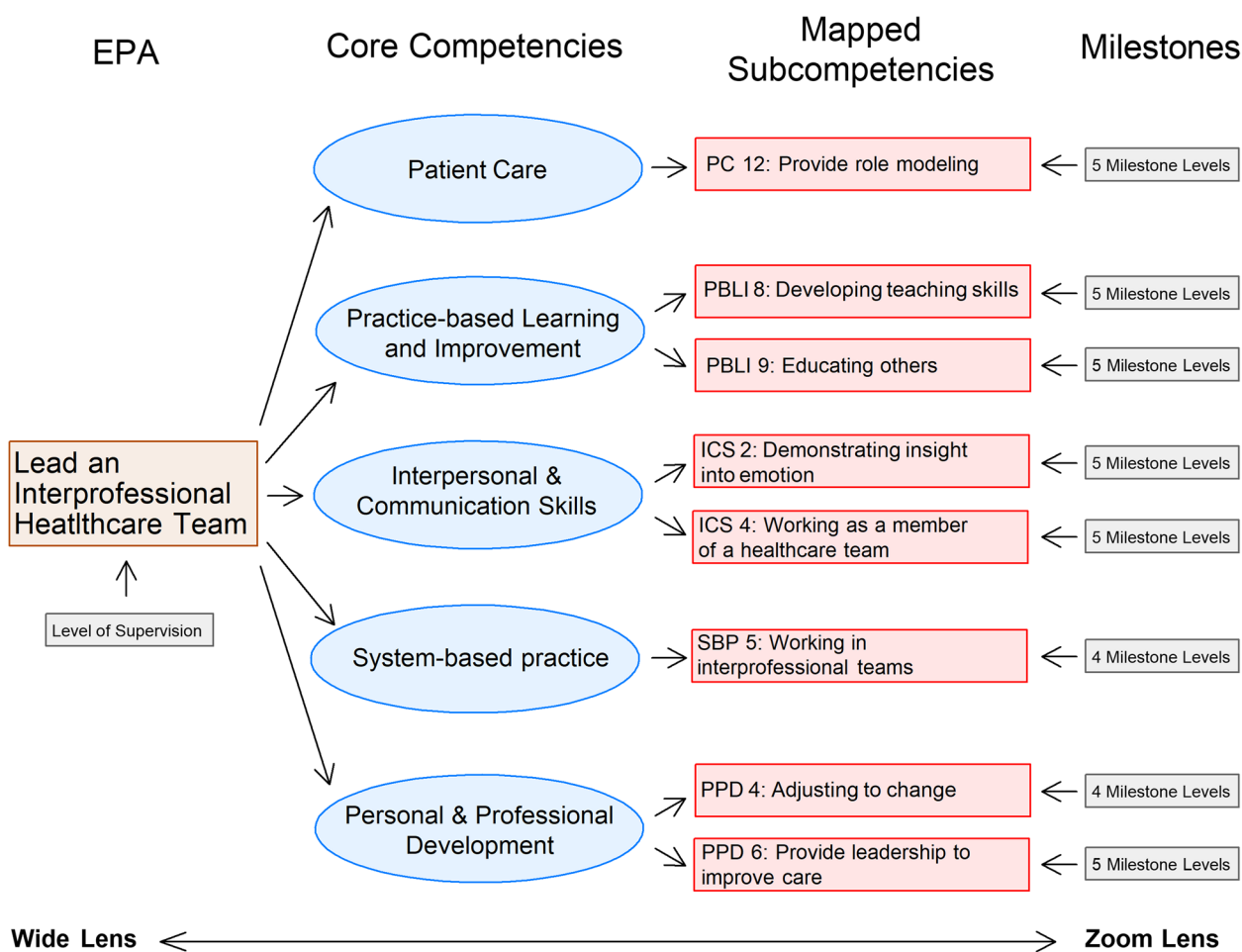
Fig. 1 Schematic showing the relationship of the Lead an Interprofessional Healthcare Team EPA with mapped core competencies and subcompetencies. Eight subcompetencies map to this EPA. Milestones were created for all subcompetencies but of the 8, only 5 are reported to the Accreditation Council for Graduate Medical Education. Personal & Professional Development (PPD) is a core competency unique to Pediatrics. Abbreviations: PC = patient care, PBLI = practice-based learning and improvement, ICS = interpersonal and communication skills, SBP = system-based practice

One week before the Clinical Competency Committee (CCC) meeting, we asked fellowship program directors (FPDs) to assign a LOS rating for each fellow for 6 of the 7 EPAs common to all pediatric subspecialties (Common Pediatric Subspecialty EPAs; Table 1) [22]. Then, at the CCC meeting, we asked the members to first assign a milestone level to all 29 subcompetencies mapped to these six EPAs. Of note, in Pediatrics, all subspecialties utilize the same milestones. CCCs then assigned a LOS rating for each fellow for each Common Pediatric Subspecialty EPA. We provided no specific instructions to the FPD or CCC members about the procedure to determine fellow ratings or faculty development about EPAs or the EPA LOS scales. Representatives of the 14 subspecialties contributed to the development of the EPA LOS scales and the validity evidence for them has previously been published [23, 24]. Designed to be intuitive, these ordinal 5-level scales are based upon direct, indirect and no supervision with case complexity being a variable in determining the need for supervision at some levels for some EPAs (Table 1).

The anonymity of trainees was ensured by creating a unique participant identifier number using an algorithm developed by APPD LEARN [25]. Once this ID was created, we provided specific links to the online data collection instruments. In the survey instrument, we first elicited milestone ratings for each of the subcompetencies grouped by the 6 core competencies and then obtained LOS ratings for each EPA. When presenting the subcompetency, we displayed the subcompetency name and descriptions for each milestone; when presenting each EPA, we displayed the title of the EPA and the associated functions necessary to carry out the activities followed by the LOS scale.

Mink *et al. BMC Medical Education*   (2023) 23:720

Page 5 of 14

We also collected information about each fellow's sub-specialty and year of fellowship, institution, the number of fellows in the program, how long the FPD served in this role, and FPDs self-reported understanding of EPAs (unfamiliar, basic, in-depth, or expert). We also asked whether the FPD was a member of the CCC since FPD participation on the CCC may influence assignment of trainee ratings [24]. Details about the data collection tools have been previously described [23]. We collected data in fall 2014 and spring 2015. The abbreviations for each EPA are listed in Table 1.

For each EPA, we computed an unweighted composite milestone score by averaging the milestone levels for the subcompetencies mapped to that EPA. We compared LOS ratings and unweighted composite score for trainees at each data collection period using linear mixed models adjusting for repeated measures and clustering within programs.

We computed a weighted composite milestone score for each EPA by using a confirmatory factor analysis procedure to fit path coefficients and mean structures between each EPA's LOS and its mapped subcompetencies, adjusting for clustering in program, and then used the path coefficients to generate a weighted average of the milestone levels. To assess the fit of the procedure, we examined the comparative fit index and the root mean squared error of approximation using the entire sample; to guard against overfitting, we also conducted a fivefold cross-validation bootstrap process, fitting the path coefficients on 80% of the data and making predictions on the remaining 20%, repeating the process for each fold and averaging 500 replications.

We tested the hypothesis that composite milestone scores would be correlated with LOS ratings using Spearman's ρ. We tested the hypothesis that weighted composites would outperform unweighted composites by comparing confidence intervals around the ρ values for the weighted and unweighted composites, and similarly tested differences between unweighted composites at programs where the FPD did or did not serve on the CCC. We tested the hypothesis that using all critical subcompetencies would better predict LOS than using only ACGME-reported milestones in a similar fashion, and directly compared the fit of the nested weighted confirmatory factor analysis models using a likelihood ratio χ^2 test.

We generated equations to predict milestone levels using the path coefficients in the confirmatory factor analysis for each model. For external validation of the predicted levels, we used spring 2019 EPA LOS ratings that were obtained in a recently completed study [26]. EPA LOS ratings were collected in the same manner as described above except that milestone levels were

not obtained. Spring 2019 milestone levels were provided by the ACGME through a data sharing agreement with APPD LEARN. With these data, we examined the goodness-of-fit using the ACGME equations.

Using the model with the best fit and parsimony, we constructed receiver operating characteristic (ROC) curves for the ability of that model's composite milestone score to discriminate between decisions affirming or refuting entrustment, using levels 4 or 5 as the minimum level for affirmation of entrustment. Data analyses were conducted using R 3.6 (R Core Team, Vienna, Austria).

Results

We assessed 1040 fellows in fall and 1048 in spring, representing about 27% of all pediatric fellows [27] Data were submitted from 78 and 82 different institutions and 209 and 212 programs in fall and spring, respectively. In both periods, 79% (11/14) of subspecialties met our goal of having at least 20% of their subspecialty programs provide data. FPDs were a member of the CCC for 57.5% (598/1040) of ratings in the first data collection period and 55.7% (584/1048) in the second. FPDs completed their assessments a median of 6[IQR 1–9] and 7 [1-11] days before the CCC meeting in fall and spring, respectively.

Mean EPA LOS and unweighted composite milestone scores for each period are displayed in Table 2. Both EPA LOS and milestone score increased from the fall to the spring ($p < 0.001$ for each EPA, adjusting for repeated measures and clustering of trainees in programs).

Table 2 Overall mean (95% CI) unweighted composite subcompetency milestone scores and EPA level of supervision ratings in fall 2014 and spring 2015

EPA	Period	Unweighted composite milestone score[a]	Mean EPA LOS Rating[a]
QI	Fall	3.1 (3.1–3.1)	2.7 (2.6–2.7)
	Spring	3.3 (3.3–3.4)	3.0 (2.9–3.0)
Consultation	Fall	3.3 (3.3–3.3)	3.4 (3.4–3.5)
	Spring	3.5 (3.5–3.6)	3.8 (3.7–3.8)
Management	Fall	3.2 (3.2–3.2)	2.8 (2.7–2.8)
	Spring	3.5 (3.5–3.5)	3.2 (3.1–3.2)
Handover	Fall	3.3 (3.3–3.4)	3.7 (3.6–3.7)
	Spring	3.6 (3.6–3.6)	4.0 (4.0–4.1)
Leadteam	Fall	3.3 (3.2–3.3)	3.0 (3.0–3.0)
	Spring	3.5 (3.5–3.5)	3.4 (3.3–3.4)
Leadprof	Fall	3.2 (3.1–3.2)	2.7 (2.6–2.7)
	Spring	3.4 (3.4–3.4)	3.0 (3.0–3.1)

Abbreviations: *EPA* Entrustable professional activities, *LOS* level of supervision

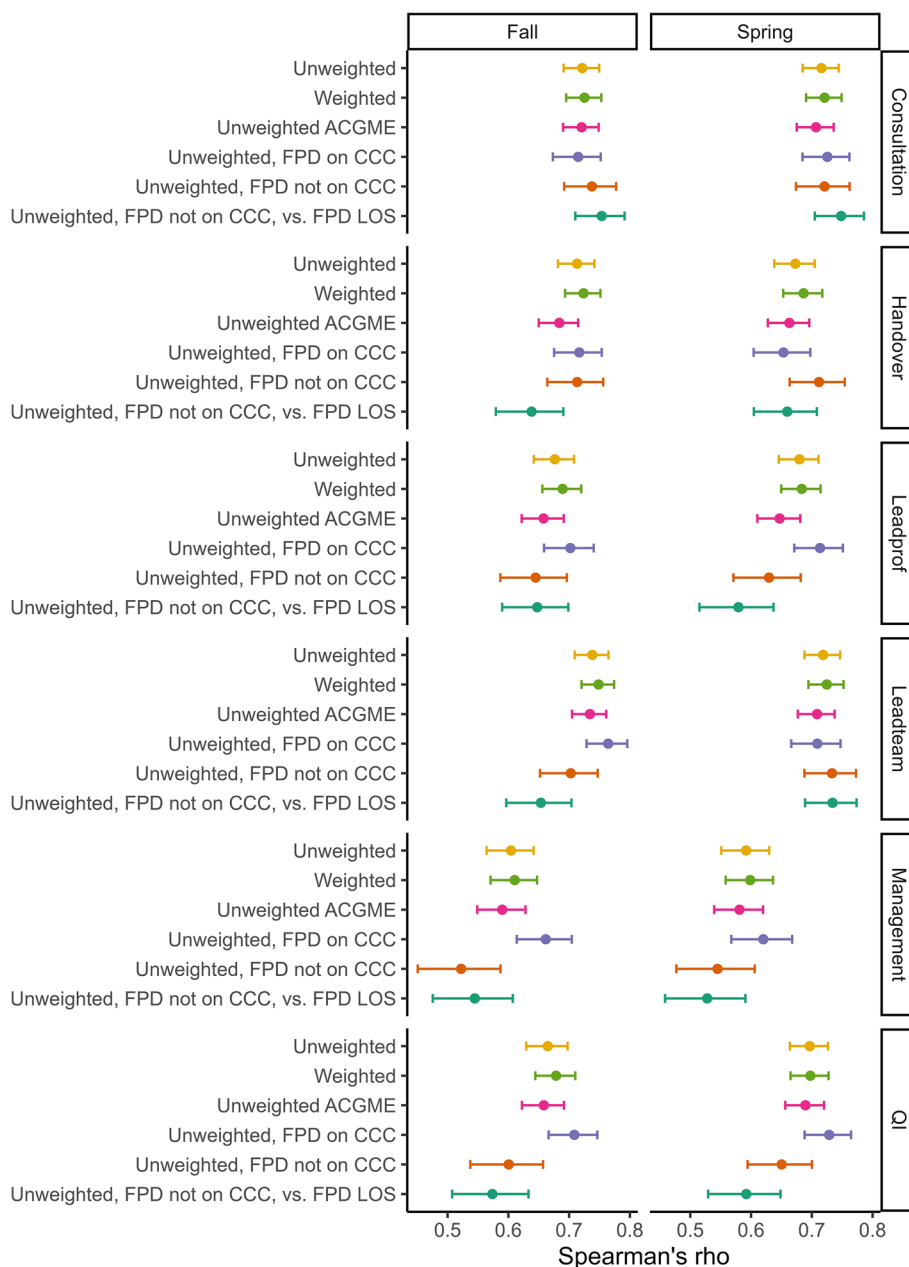[a] $p < 0.05$ fall vs. spring for all EPAs

Fig. 2 Graph showing Spearman Rho correlations (95% confidence intervals) of EPA level of supervision ratings by the Clinical Competency Committee with unweighted (#1) and weighted (#2) composite score using all mapped milestones, unweighted using only the ACGME reported milestones (#3) and unweighted using milestones from when the fellowship program director was (#4) or was not (#5) a member of the Clinical Competency Committee. The last graph in each group (#6) shows the correlation of EPA level of supervision ratings made independently by the fellowship program director with milestones for when the program director was not a member of the clinical competency committee. Data are from the fall 2014 and spring 2015. Abbreviations: ACGME = Accreditation Council for Graduate Medical Education, CCC = Clinical Competency Committee, EPA = Entrustable professional activities, FPD = fellowship program director; LOS = level of supervision

Testing hypothesis 1: There is a strong correlation between EPA LOS rating with the average score of the mapped subcompetency milestone levels needed to perform the EPA

There was moderate to strong correlation between the unweighted composite milestone score and EPA LOS,

ranging from 0.59 for the Management EPA to 0.74 for Leadteam, supporting our first hypothesis (Fig. 2, Table 3). There was no difference in the correlations between the two periods ($p > 0.05$). Correlations between LOS ratings made independently by the FPD before the CCC meeting with composite milestone scores

Mink *et al. BMC Medical Education*   (2023) 23:720

Page 7 of 14

Table 3 Correlation [Rho (95% CI)] of the unweighted and weighted composite subcompetency milestone score with EPA level of supervision ratings in fall 2014 and spring 2015 using all milestone data, only ACGME reported milestones and whether the program director was or was not a member of Clinical Competency Committee

Source of Data		All CCC members	All CCC members	All CCC members, only ACGME reported milestones	FPD on CCC	FPD not on CCC	FPD not on CCC
EPA	Period	Unweighted composite milestone score vs CCC LOS[a]	Weighted composite milestone score vs CCC LOS[a]	Unweighted composite milestone score vs CCC LOS[a]	Unweighted composite milestone score vs CCC LOS[a]	Unweighted composite milestone score vs CCC LOS[a]	Unweighted composite milestone score vs FPD LOS[a]
QI	Fall	0.66(0.63—0.70)	0.68(0.64–0.71)	0.66(0.62–0.69)	0.71(0.67–0.75)	0.60(0.54–0.66)[b]	0.57(0.51–0.63)[c]
	Spring	0.70(0.66—0.73)	0.70(0.67–0.73)	0.69(0.66–0.72)	0.73(0.69–0.76)	0.65(0.59–0.70)	0.59(0.53–0.65)[c]
Consultation	Fall	0.72(0.69—0.75)	0.73(0.70–0.75)	0.72(0.69–0.75)	0.71(0.67–0.75)	0.74(0.69–0.78)	0.75(0.71–0.79)
	Spring	0.72(0.69—0.74)	0.72(0.69–0.75)	0.71(0.68–0.74)	0.73(0.68–0.76)	0.72(0.67–0.76)	0.75(0.70–0.79)
Management	Fall	0.60(0.56—0.64)	0.61(0.57–0.65)	0.59(0.55–0.63)	0.66(0.61–0.70)	0.52(0.45–0.59)[b]	0.54(0.48–0.61)
	Spring	0.59(0.55—0.63)	0.60(0.56–0.64)	0.58(0.54–0.62)	0.62(0.57–0.67)	0.54(0.47–0.61)	0.53(0.46–0.59)
Handover	Fall	0.71(0.68—0.74)	0.72(0.69–0.75)	0.68(0.65–0.71)	0.72(0.68–0.75)	0.71(0.66–0.76)	0.64(0.58–0.69)
	Spring	0.67(0.64—0.71)	0.69(0.65–0.72)	0.66(0.63–0.70)	0.65(0.60–0.70)	0.71(0.66–0.75)	0.66(0.60–0.71)
Leadteam	Fall	0.74(0.71—0.76)	0.75(0.72–0.77)	0.73(0.70–0.76)	0.76(0.73–0.80)	0.70(0.65–0.75)	0.65(0.60–0.70)[c]
	Spring	0.72(0.68—0.75)	0.72(0.69–0.75)	0.71(0.68–0.74)	0.71(0.67–0.75)	0.73(0.69–0.77)	0.73(0.69–0.77)
Leadprof	Fall	0.68(0.64—0.71)	0.69(0.66–0.72)	0.66(0.62–0.69)	0.70(0.66–0.74)	0.64(0.59–0.70)	0.65(0.59–0.70)
	Spring	0.68(0.64—0.71)	0.68(0.65–0.71)	0.65(0.61–0.68)	0.71(0.67–0.75)	0.63(0.57–0.68)	0.58(0.52–0.64)[c]

[a] all values $p < 0.001$; [b] $p < 0.05$ compared with unweighted composite milestone score for programs in which the FPD is on the CCC; [c] $p < 0.05$ compared with unweighted composite milestone score with all members on CCC

ACGME Accreditation Council for Graduate Medical Education, *CCC* Clinical Competency Committee, *FPD* program director, *LOS* level of supervision

were similar to those made by the CCC when the FPD was not a member. In addition, when examining the associations in programs where the FPD was not a CCC member versus where the FPD was a member, the correlations were somewhat lower for some EPAs (QI and Management in the fall). Otherwise, they were not significantly different. The significant associations between milestone and EPA LOS ratings persisted after adjustment for institution, subspecialty, and program and FPD characteristics.

Testing hypothesis 2: Some subcompetencies would be more critical than others such that weighted scores will have a stronger correlation

Correlations calculated using a weighted composite milestone score were not significantly better than those calculated using an unweighted score, counter to our second hypothesis. The most parsimonious best-fitting model was thus the unweighted ACGME-reported-milestones-only composite score. Figure 3 shows ROC curves using unweighted ACGME-reported-milestones-only composite score from the spring for the 6 EPAs to predict entrustment based upon a minimum EPA LOS of 4 or 5. The area under the curve (AUC) was excellent, ranging from 0.81 (95% CI: 0.78–0.84) for Management to 0.90 (0.86–0.94) for QI. When assuming entrustment

based upon attaining a LOS of 5, the AUCs were similar to those using a minimum of level 4 ($p > 0.05$). For each EPA, there was no difference between the AUC in fall and spring ($p > 0.05$) or based upon FPD CCC membership ($p > 0.05$).

Testing hypothesis 3: If only those milestones required for reporting to the ACGME were included in the analysis, the correlation between EPA LOS rating and average milestone level would be weaker

Goodness-of-fit for weighted models using all milestones or only those reported to the ACGME were both excellent and did not differ ($p = 0.72$), counter to our third hypothesis that the correlation using all mapped milestones would be stronger. The comparative fit index and root mean square error of approximation of models using all milestones were 0.999 (> 0.95 is excellent) and 0.034 (< 0.05 is excellent), respectively, while values using only the ACGME reported milestones were 0.998 and 0.043 [28]. Prediction equations for each model are shown in Table 4.

The external validation sample included 1373 EPA LOS ratings from 503 (36.6%) first year, 448 (32.6%) second year and 422 (30.7%) third year fellows. The comparative fit index and root mean square error of approximation of models using the ACGME prediction equations were
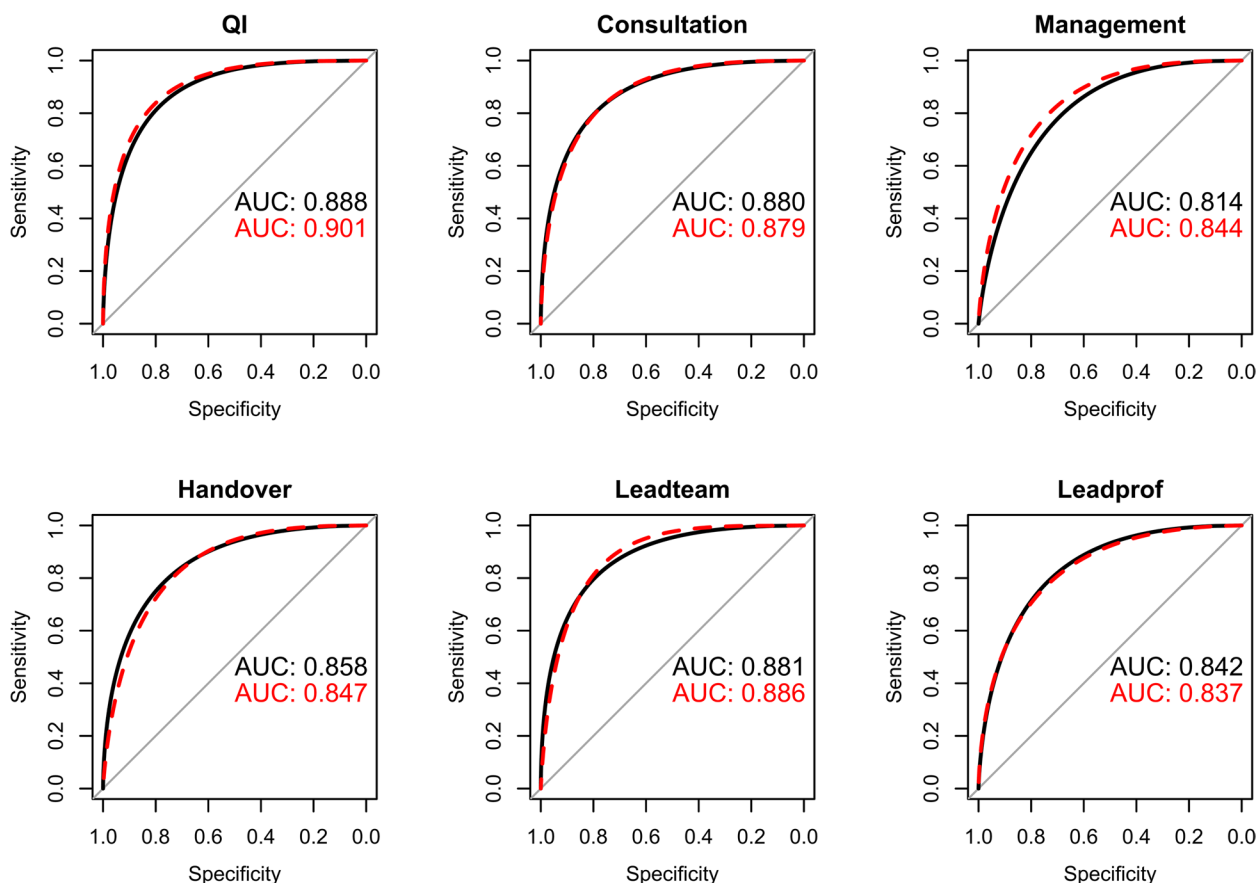
Mink *et al. BMC Medical Education*   (2023) 23:720

Page 8 of 14

Fig. 3 ROC curves and area under the curve for spring 2015 data using unweighted Accreditation Council for Graduate Medical Education subcompetency milestone composite score to achieve EPA level of supervision ratings of 4 or 5 (solid line; black) or only 5 (dashed line; red). Ratings utilized data from all members of the Clinical Competency Committee. Area under the curve in black is based upon a rating of 4 or 5 while that in red used a rating of 5. Abbreviations: ROC=receive operating characteristics, AUC=area under the curve, EPA=Entrustable professional activities

0.994 (>0.95 is excellent) and 0.071 (0.06–0.08 is fair), respectively [28].

Discussion

In support of our first hypothesis, we found a strong relationship between milestone levels for subcompetencies mapped to an EPA and the LOS ratings for that EPA, providing validity evidence for both approaches. Our data do not support the second hypothesis in that we found the relationship between EPA LOS and milestone scores was nearly identical whether we used the unweighted or weighted milestone scores. Likewise, the relationship between milestone level and EPA LOS rating was similar when only the ACGME reported milestones were utilized in the model compared with using milestones from all 29 subcompetencies mapped to the six EPAs.

Our results are similar to those of Larrabee et al., who examined the association between 27 EPAs that they developed for 4 core rotations in pediatric residency with milestones mapped to these EPAs [29]. They found a strong correlation between the two ratings, with an overall median R^2 of 0.81. Although these investigators focused on residents and used a different LOS scale, the concordance between Larrabee's findings and ours nevertheless suggests that the relationship between milestones and EPA LOS is generalizable and not dependent upon a particular group of trainees or a specific LOS scale.

The areas under the curve for all EPAs were very high. Except for one, this was irrespective of whether entrustment was set at level 5 or at level 4 or 5, indicating that the relationship was not solely dependent upon how entrustment was defined. While executing the EPA without supervision is the goal, not all FPDs believe that fellows must achieve level 5 (unsupervised practice) in all EPAs to graduate, indicating that some supervision may still be needed [30–32]. Also, for some EPAs, while the correlations were somewhat weaker if the FPD was not a member on the CCC, the differences were small and the AUCs for the ROC curves were not affected.

Mink *et al. BMC Medical Education*   (2023) 23:720

Page 9 of 14

Table 4 Equations used to determine subcompetency milestone level based upon EPA level of supervision ratings and whether all mapped subcompetency milestones were used in the model or only those reported to the ACGME[a]

Subcompetency	Utilizing all 29 Subcompetency Milestones	Utilizing only 21 ACGME Reported Subcompetency Milestones
ICS1	$1.1699 + 0.6598(LOS_{Consultation})$	-
ICS2	$1.7798 + 0.5334(LOS_{Leadteam})$	-
ICS3	$1.1496 + 0.1667(LOS_{Leadprof}) + 0.4791(LOS_{Handover})$	$1.2272 + 0.2205(LOS_{Leadprof}) + 0.4192(LOS_{Handover})$
ICS4	$1.724 - 0.0098(LOS_{Leadprof}) + 0.5871(LOS_{Leadteam}) - 0.0277(LOS_{Consultation})$	$1.5472 + 0.0473(LOS_{Leadprof}) + 0.3620(LOS_{Leadteam}) + 0.1742(LOS_{Consultation})$
ICS5	$1.0062 + 0.6329(LOS_{Consultation})$	$1.0049 + 0.63332(LOS_{Consultation})$
ICS6	$1.4236 + 0.2803(LOS_{Management} + 0.291(LOS_{Handover})$	-
MK2	$1.1398 + 0.0378(LOS_{Leadprof}) + 0.6143(LOS_{Consultation})$	$1.1530 + 0.0568(LOS_{Leadprof}) + 0.5954(LOS_{Consultation})$
PBLI1	$1.3572 + 0.605(LOS_{Consultation})$	$1.3563 + 0.6052(LOS_{Consultation})$
PBLI4	$1.4387 + 0.636(LOS_{QI})$	$1.4295 + 0.6391(LOS_{QI})$
PBLI5	$1.0036 + 0.6473(LOS_{Handover})$	-
PBLI7	$1.7149 + 0.3266(LOS_{QI}) + 0.2229(LOS_{Handover})$	$1.6741 + 0.2919(LOS_{QI}) + 0.2586(LOS_{Handover})$
PBLI8	$1.5107 + 0.604(LOS_{Leadteam})$	-
PBLI9	$1.6212 + 0.5898(LOS_{Leadteam})$	$1.6250 + 0.5887(LOS_{Leadteam})$
PC12	$1.624 - 0.01(LOS_{Leadprof}) + 0.5971(LOS_{Leadteam})$	$1.6256 + 0.0482(LOS_{Leadprof}) + 0.5441(LOS_{Leadteam})$
PC2	$0.8874 + 0.7066(LOS_{Consultation})$	-
PC3	$0.9502 + 0.6875(LOS_{Handover})$	$0.9803 + 0.6800(LOS_{Handover})$
PC6	$1.1439 + 0.6207(LOS_{Consultation})$	$1.1426 + 0.6210(LOS_{Consultation})$
PPD4	$1.8176 + 0.0532(LOS_{Leadprof}) + 0.3391(LOS_{Leadteam})$	-
PPD6	$1.4679 - 0.0286(LOS_{Leadprof}) + 0.6234(LOS_{Leadteam})$	$1.1426 - 0.0272(LOS_{Leadprof}) + 0.5724(LOS_{Leadteam})$
PPD8	$1.1983 + 0.6201(LOS_{Consultation})$	$1.1972 + 0.6204(LOS_{Consultation})$
PROF2	$2.0305 + 0.5323(LOS_{Management})$	$2.0249 + 0.5341(LOS_{Management})$
PROF4	$2.3314 + 0.3266(LOS_{QI})$	-
SBP1	$1.6532 + 0.5923(LOS_{Management})$	$1.6472 + 0.5941(LOS_{Management})$
SBP2	$1.3604 + 0.5652(LOS_{Consultation})$	$1.3596 + 0.5693(LOS_{Consultation})$
SBP3	$1.4241 + 0.2285(LOS_{Management}) + 0.381(LOS_{QI})$	$1.4136 + 0.3005(LOS_{Management}) + 0.3082(LOS_{QI})$
SBP4	$1.501 + 0.0874(LOS_{Leadprof}) + 0.5383(LOS_{QI})$	-
SBP5	$2.0918 + 0.021(LOS_{Management}) + 0.3295(LOS_{Leadteam}) + 0.0571(LOS_{QI})$	$2.0879 + 0.0364(LOS_{Management}) + 0.2791(LOS_{Leadteam}) + 0.0994(LOS_{QI})$
SBP6	$1.7567 + 0.1286(LOS_{Managment}) + 0.4518(LOS_{QI})$	$1.7454 + 0.2224(LOS_{Management}) + 0.3562(LOS_{QI})$
SBP7	$1.6227 + 0.095(LOS_{Leadprof}) + 0.3355(LOS_{QI})$	-

ACGME Accreditation Council for Graduate Medical Education, *LOS* level of supervision

[a] Subcompetencies are those associated with the Pediatric Milestones; [42] subscripts indicate specific EPA

We constructed equations to predict milestone level based upon EPA LOS rating that had an excellent goodness of fit. In both the derivation and validation samples, the comparative fit index was excellent. While there was a slight decline in the root mean square error of approximation using the data 4 years later, coupled with the strong comparative fix index, the overall goodness-of-fit was still very good. Since the spring 2019 data represent assessments of different fellows, and likely include CCCs that had differing compositions, this shows that the equations maintained their precision over time.

Showing that there is a strong relationship between milestones and EPA LOS helps to address FPD's concerns about the additional work involved as EPAs become a required element of trainee assessment [33–35]. Faculty find using EPA LOS scales very intuitive and generating milestone levels based on EPA LOS ratings should be timesaving [33]. These predicted milestone ratings can serve as a starting point when the CCC discusses each trainee and makes the final assignments. We used milestones version 1.0 in our study to develop the equations, but milestones 2.0 will shortly be implemented across all specialties and subspecialties. Our finding that models using all milestones compared with only those

reported to the ACGME are similar will make it easier to revise the equations with updated milestones.

We found little difference in the correlations between EPA LOS and milestones based on whether or not the FPD was a member of the CCC. These findings are consistent with our previous report that the association between FPD and CCC assignment of EPA LOS is strong [24]. With the exception of the Management EPA in the fall, the correlations for when the FPD made the assessments independently of the CCC were also similar. It is reassuring that the relationship between LOS and milestones is not affected by FPD membership on the CCC.

While both approaches to assessment are highly related, it is important to recognize the contribution of each in creating an overall program of trainee assessment [17, 36–38]. As program directors and CCCs make decisions about trainee progression toward unsupervised practice, the need to focus more on either EPA LOS ratings or milestone levels may depend on the circumstances. For high performing trainees who require minimal supervision to execute an EPA, milestone levels may be less important than for a trainee who is early in development and requires more supervision. In the latter case, the descriptive language of the skills included in each milestone level can help the trainee focus on improvements needed to effectively perform the EPA, especially if they are below the normative national standards or not meeting program expectations [39–41].

There are several limitations to this study. We asked FPDs to assign milestones before assigning LOS for the EPAs rather than randomizing the assessments. This could have biased their rating for EPA LOS. The initial data collection period was the first time FPDs had to report milestones to the ACGME and assign EPA LOS ratings. With more experience, the application of these assessments may change, although we saw no difference in results between the two reporting periods. In addition, the goodness-of-fit of the equations using the ACGME-reported milestones suggests that there has not been much change in in the milestone-LOS relationship over time. Finally, we used the Common Pediatric Subspecialty EPAs, and the findings cannot necessarily be extrapolated to assessments made using EPAs developed by other specialties.

Conclusions

We found strong agreement between assessments based on subcompetency milestones with those using EPA LOS but no difference if the subcompetencies were weighted or if only the ACGME reported subcompetencies were used. In addition, these data provide additional validity evidence for both types of assessment. We were also able to develop equations to generate milestone levels based on EPA supervision ratings using the ACGME reported subcompetencies. This will help to address the time burden faced by educators while also allowing them the flexibility to use EPAs and milestones as appropriate in assessing their trainees and "developing "a program of assessment fit for purpose" [17].

Abbreviations

ACGME	Accreditation Council for Graduate Medical Education
APPD LEARN	Association of Pediatric Program Directors Longitudinal Educational Assessment Research Network
AUC	Area under the curve
CCC	Clinical Competency Committee
Consultation EPA	Provide consultation to other healthcare providers caring for children and adolescents and refer patients requiring further consultation to other subspecialty providers if necessary
EPA	Entrustable Professional Activities
FPD	Fellowship program director
Handover	EPA: Facilitate handovers to another healthcare provider either within or across settings
ICS	Core competency: Interpersonal and communication skills
Leadteam	EPA: Lead an interprofessional health care team
Leadprof	EPA: Lead within the subspecialty profession
LOS	Level of supervision
Management	EPA: Contribute to the fiscally sound, equitable and collaborative management of a healthcare workplace
PBLI	Core competency: Practice-based learning and improvement
PC	Core competency: Patient care
PPD	Core competency in Pediatrics: Personal and Professional Development
QI	EPA: Apply public health principles and quality improvement methods to improve population health
ROC	Receiver operating characteristic
SBP	Core competency: System-based practice

Elizabeth S Jacobs, Brown University/Rhode Island Hospital-Lifespan

Jennifer Welch, Brown University/Rhode Island Hospital-Lifespan

Deepak Kumar, Case Western Reserve University/Metro Health

Katherine Mason, Case Western Reserve University/Rainbow Babies and Children's Hospital

Nancy Roizen, Case Western Reserve University/Rainbow Babies and Children's Hospital

Jerri A. Rose, Case Western Reserve University/Rainbow Babies and Children's Hospital

Brooke Bokor, Children's National Medical Center/George Washington University

Jennifer I Chapman, Children's National Medical Center/George Washington University

Lowell Frank, Children's National Medical Center/George Washington University

Iman Sami, Children's National Medical Center/George Washington University

Jennifer Schuette, Children's National Medical Center/George Washington University

Ramona E Lutes, Children's Hospital Medical Center of Akron

Stephanie Savelli, Children's Hospital Medical Center of Akron

Rambod Amirnovin, Children's Hospital of Los Angeles

Rula Harb, Children's Hospital of Los Angeles

Roberta Kato, Children's Hospital of Los Angeles

Karen Marzan, Children's Hospital of Los Angeles

Roshanak Monzavi, Children's Hospital of Los Angeles

Doug Vanderbilt, Children's Hospital of Los Angeles

Lesley Doughty, Cincinnati Children's Hospital Medical Center

Constance McAneney, Cincinnati Children's Hospital Medical Center

Ward Rice, Cincinnati Children's Hospital Medical Center

Lea Widdice, Cincinnati Children's Hospital Medical Center

Fran Erenberg, Cleveland Clinic Children's Hospital

Blanca E Gonzalez, Cleveland Clinic Children's Hospital

Deanna Adkins, Duke University Medical Center

Deanna Green, Duke University Medical Center

Aditee Narayan, Duke University Medical Center

Kyle Rehder, Duke University Medical Center

Joel Clingenpeel, Eastern Virginia Medical School

Suzanne Starling, Eastern Virginia Medical School

Heidi Eigenrauch Karpen, Emory University School of Medicine

Kelly Rouster-Stevens, Emory University School of Medicine

Jatinder Bhatia, Georgia Regents University/Medical College of Georgia

John Fuqua, Indiana University Medical Center-Riley Hospital for Children

Jennifer Anders, Johns Hopkins University

Maria Trent, Johns Hopkins University

Rangasamy Ramanathan, LAC+USC Medical Center

Yona Nicolau, Loma Linda University Children's Hospital

Allen J. Dozor, Maria Fareri Children's Hospital Westchester Medical Center/New York Medical College

Thomas Bernard Kinane, Massachusetts General Hospital

Takara Stanley, Massachusetts General Hospital

Amulya Nageswara Rao, Mayo Clinic College of Medicine (Rochester)

Meredith Bone, McGaw Medical Center of Northwestern University/Lurie Children's Hospital of Chicago

Lauren Camarda, McGaw Medical Center of Northwestern University/Lurie Children's Hospital of Chicago

Viday Heffner, Medical College of Wisconsin, Children's Hospital of Wisconsin

Olivia Kim, Medical College of Wisconsin, Children's Hospital of Wisconsin

Jay Nocton, Medical College of Wisconsin, Children's Hospital of Wisconsin

Angela L Rabbitt, Medical College of Wisconsin, Children's Hospital of Wisconsin

Richard Tower, Medical College of Wisconsin, Children's Hospital of Wisconsin

Michelle Amaya, Medical University of South Carolina

Jennifer Jaroscak, Medical University of South Carolina

James Kiger, Medical University of South Carolina

Michelle Macias, Medical University of South Carolina

Olivia Titus, Medical University of South Carolina

Modupe Awonuga, Michigan State University

Karen Vogt, National Capital Consortium (Walter Reed)

Anne Warwick, National Capital Consortium (Walter Reed)

Dan Coury, Nationwide Children's Hospital/Ohio State University

Mark Hall, Nationwide Children's Hospital/Ohio State University

Megan Letson, Nationwide Children's Hospital/Ohio State University

Melissa Rose, Nationwide Children's Hospital/Ohio State University

Julie Glickstein, New York Presbyterian Hospital-Columbia campus/Morgan Stanley Children's Hospital

Sarah Lusman, New York Presbyterian Hospital-Columbia campus/Morgan Stanley Children's Hospital

Cindy Roskind, New York Presbyterian Hospital-Columbia campus/Morgan Stanley Children's Hospital

Karen Soren, New York Presbyterian Hospital-Columbia campus/Morgan Stanley Children's Hospital

Jason Katz, Nicklaus Children's Hospital-Miami Children's Hospital

Lorena Siqueira, Nicklaus Children's Hospital-Miami Children's Hospital

Mark Atlas, North Shore-LIJ/ Cohen Children's Medical Center of New York

Andrew Blaufox, North Shore-LIJ/ Cohen Children's Medical Center of New York

Beth Gottlieb, North Shore-LIJ/ Cohen Children's Medical Center of New York

David Meryash, North Shore-LIJ/Cohen Children's Medical Center of New York

Patricia Vuguin, North Shore-LIJ/Cohen Children's Medical Center of New York

Toba Weinstein, North Shore-LIJ/Cohen Children's Medical Center of New York

Laurie Armsby, Oregon Health and Science University Hospital

Lisa Madison, Oregon Health and Science University Hospital

Brian Scottoline, Oregon Health and Science University Hospital

Evan Shereck, Oregon Health and Science University Hospital

Michael Henry, Phoenix Children's Hospital

Patricia A. Teaford, Phoenix Children's Hospital

Sarah Long, St. Christopher's Hospital for Children

Laurie Varlotta, St. Christopher's Hospital for Children

Alan Zubrow, St. Christopher's Hospital for Children

Courtenay Barlow, Stanford University/Lucile Packard Children's Hospital

Heidi Feldman, Stanford University/Lucile Packard Children's Hospital

Hayley Ganz, Stanford University/Lucile Packard Children's Hospital

Paul Grimm, Stanford University/Lucile Packard Children's Hospital

Tzielan Lee, Stanford University/Lucile Packard Children's Hospital

Leonard B. Weiner, SUNY Upstate Medical University

Zarela Molle-Rios, Thomas Jefferson University Hospital/duPont Hospital for Children Program/Christiana

Nicholas Slamon, Thomas Jefferson University Hospital/duPont Hospital for Children Program/Christiana

Ursula Guillen, Thomas Jefferson University Hospital/duPont Hospital for Children Program/Christiana

Karen Miller, Tufts Medical Center

Myke Federman, UCLA Medical Center/Mattel Children's Hospital

Randy Cron, University of Alabama Medical Center at Birmingham

Wyn Hoover, University of Alabama Medical Center at Birmingham

Tina Simpson, University of Alabama Medical Center at Birmingham

Margaret Winkler, University of Alabama Medical Center at Birmingham

Nada Harik, University of Arkansas for Medical Sciences

Ashley Ross, University of Arkansas for Medical Sciences

Omar Al-Ibrahim, University of Buffalo-Women and Children's Hospital of Buffalo

Frank P. Carnevale, University of Buffalo-Women and Children's Hospital of Buffalo

Wayne Waz, University of Buffalo-Women and Children's Hospital of Buffalo

Fayez Bany-Mohammed, University of California Irvine/Miller Children's Hospital

Jae H. Kim, University of California San Diego Medical Center/Rady Children's Hospital

Beth Printz, University of California San Diego Medical Center/Rady Children's Hospital

Mike Brook, University of California San Francisco

Michelle Hermiston, University of California San Francisco

Erica Lawson, University of California San Francisco

Sandrijn van Schaik, University of California San Francisco

Alisa McQueen, University of Chicago Medical Center

Karin Vander Ploeg Booth, University of Chicago Medical Center

Melissa Tesher, University of Chicago Medical Center

Jennifer Barker, University of Colorado School of Medicine/ Children's Hospital Colorado

Sandra Friedman, University of Colorado School of Medicine/ Children's Hospital Colorado

Ricky Mohon, University of Colorado School of Medicine/ Children's Hospital Colorado

Mink *et al. BMC Medical Education*   (2023) 23:720

Page 12 of 14

Andrew Sirotnak, University of Colorado School of Medicine/ Children's Hospital Colorado

John Brancato, University of Connecticut/Connecticut Children's Medical Center

Wael N. Sayej, University of Connecticut/Connecticut Children's Medical Center

Nizar Maraqa, University of Florida College of Medicine at Jacksonville

Michael Haller, University of Florida College of Medicine-J Hillis Miller Health Center

Brenda Stryjewski, University of Hawaii/Tripler Army Medical Center

Pat Brophy, University of Iowa Hospitals and Clinics

Riad Rahhal, University of Iowa Hospitals and Clinics

Ben Reinking, University of Iowa Hospitals and Clinics

Paige Volk, University of Iowa Hospitals and Clinics

Kristina Bryant, University of Louisville

Melissa Currie, University of Louisville

Katherine Potter, University of Louisville

Alison Falck, University of Maryland School of Medicine

Joel Weiner, University of Massachusetts Memorial Medical Center

Michele M. Carney, University of Michigan Medical Center

Barbara Felt, University of Michigan Medical Center

Andy Barnes, University of Minnesota Medical Center

Catherine M Bendel, University of Minnesota Medical Center

Bryce Binstadt, University of Minnesota Medical Center

Karina Carlson, University of Missouri at Kansas City/Children's Mercy Hospital

Carol Garrison, University of Missouri at Kansas City/Children's Mercy Hospital

Mary Moffatt, University of Missouri at Kansas City/Children's Mercy Hospital

John Rosen, University of Missouri at Kansas City/Children's Mercy Hospital

Jotishna Sharma, University of Missouri at Kansas City/Children's Mercy Hospital

Kelly S. Tieves, University of Missouri at Kansas City/Children's Mercy Hospital

Hao Hsu, University of Nebraska Medical Center

John Kugler, University of Nebraska Medical Center

Kari Simonsen, University of Nebraska Medical Center

Rebecca K. Fastle, University of New Mexico School of Medicine

Doug Dannaway, University of Oklahoma Health Sciences Center

Sowmya Krishnan, University of Oklahoma Health Sciences Center

Laura McGuinn, University of Oklahoma Health Sciences Center

Mark Lowe, University of Pittsburgh Medical Center/Children's Hospital of Pittsburgh

Selma Feldman Witchel, University of Pittsburgh Medical Center/Children's Hospital of Pittsburgh

Loreta Matheo, University of Pittsburgh Medical Center/Children's Hospital of Pittsburgh

Rebecca Abell, University of Rochester School of Medicine, Golisano Children's Hospital

Mary Caserta, University of Rochester School of Medicine, Golisano Children's Hospital

Emily Nazarian, University of Rochester School of Medicine, Golisano Children's Hospital

Susan Yussman, University of Rochester School of Medicine, Golisano Children's Hospital

Alicia Diaz Thomas, University of Tennessee College of Medicine/Le Bonheur Children's Hospital

David S Hains, University of Tennessee College of Medicine/Le Bonheur Children's Hospital

Ajay J. Talati, University of Tennessee College of Medicine/Le Bonheur Children's Hospital

Elisabeth Adderson, University of Tennessee College of Medicine/St. Jude Children's Research Hospital

Nancy Kellogg, University of Texas Health Science Center School of Medicine at San Antonio

Margarita Vasquez, University of Texas Health Science Center School of Medicine at San Antonio

Coburn Allen, University of Texas Austin Dell Children's Medical School

Luc P Brion, University of Texas Southwestern Medical School

Michael Green, University of Texas Southwestern Medical School

Janna Journeycake, University of Texas Southwestern Medical School

Kenneth Yen, University of Texas Southwestern Medical School

Ray Quigley, University of Texas Southwestern Medical School

Anne Blaschke, University of Utah Medical Center

Susan L Bratton, University of Utah Medical Center

Christian Con Yost, University of Utah Medical Center

Susan P Etheridge, University of Utah Medical Center

Toni Laskey, University of Utah Medical Center

John Pohl, University of Utah Medical Center

Joyce Soprano, University of Utah Medical Center

Karen Fairchild, University of Virginia Medical Center

Vicky Norwood, University of Virginia Medical Center

Troy Alan Johnston, University of Washington/Seattle Children's Hospital

Eileen Klein, University of Washington/Seattle Children's Hospital

Matthew Kronman, University of Washington/Seattle Children's Hospital

Kabita Nanda, University of Washington/Seattle Children's Hospital

Lincoln Smith, University of Washington/Seattle Children's Hospital

David Allen, University of Wisconsin School of Medicine and Public Health

John G. Frohna, University of Wisconsin School of Medicine and Public Health

Neha Patel, University of Wisconsin School of Medicine and Public Health

Cristina Estrada, Vanderbilt University Medical Center

Geoffrey M. Fleming, Vanderbilt University Medical Center

Maria Gillam-Krakauer, Vanderbilt University Medical Center

Paul Moore, Vanderbilt University Medical Center

Joseph Chaker El-Khoury, Virginia Commonwealth University Health System

Jennifer Helderman, Wake Forest University School of Medicine

Greg Barretto, West Virginia University

Kelly Levasseur, William Beaumont Hospital

Lindsay Johnston, Yale University School of Medicine

Authors' contributions

RM conceptualized the initial study design, wrote the initial draft of the manuscript and contributed to acquisition of data, analysis and interpretation of the data and writing and editing of the manuscript. All other authors contributed to the conception and design of the study, interpretation of the data and review/editing of the manuscript. In addition, all authors approved the final manuscript and agree to be accountable for all aspects of the work.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The Institutional Review Board at the lead site, the University of Utah (#000765), gave ethics approval for the study and waived the need for informed consent on September 23, 2014. In addition, Institutional Review Board ethics approval was obtained and the need for informed consent was waived at each participating site. Research methods were in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Pediatrics, David Geffen School of Medicine at UCLA and the Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, 1124 West Carson Street, Torrance, CA 90502, USA. [2]American Board of Pediatrics, Chapel Hill, NC, USA. [3]University of Utah School of Medicine, Salt Lake City, UT, USA. [4]Department of Pediatrics, Yale School of Medicine, New Haven, CT, USA. [5]Division of Endocrinology, Department of Pediatrics, Stanford University School of Medicine, Palo Alto, CA, USA. [6]Department of Pediatrics, Duke University Medical Center, Durham, NC, USA. [7]Dana-Farber/Boston Children's Cancer and Blood Disorders Center, Boston, MA, USA. [8]Baylor College of Medicine, Houston, TX, USA. [9]Alpert Medical School of Brown University,

Mink *et al. BMC Medical Education*   (2023) 23:720

Page 13 of 14

Providence, RI, USA. [10]Developmental-Behavioral Pediatrics, Hasbro Children's Hospital, Providence, RI, USA. [11]University of Arkansas for Medical Sciences and Arkansas Children's Hospital, Little Rock, AR, USA. [12]Department of Pediatrics, University of Colorado School of Medicine, Aurora, CO, USA. [13]University of Rochester, Rochester, NY, USA. [14]Department of Pediatrics, Emory University School of Medicine and Children's Healthcare of Atlanta, Atlanta, GA, USA. [15]Division of Adolescent/Young Adult Medicine, Boston Children's Hospital, Boston, MA, USA. [16]Center for Wellbeing, Children's Mercy Hospital and University of Missouri-Kansas City School of Medicine, Kansas City, MO, USA. [17]Department of Pediatrics, Nationwide Children's Hospital and The Ohio State University College of Medicine, Columbus, OH, USA. [18]Tufts Children's Hospital, Tufts University, Boston, MA, USA. [19]University of Illinois College of Medicine at Chicago, Chicago, IL, USA.

References

1. Batalden P, Leach D, Swing S, Dreyfus H, Dreyfus S. General competencies and accreditation in graduate medical education. Health Aff. 2002;21(5):103–11.
2. Frank JR, Danoff D. The Can MEDS initiative: implementing an outcomes-based framework of physician competencies. Med Teach. 2007;29:642–7.
3. Edgar L, McLean S, Hogan S, Hamstra S, Holmboe E. The Milestones Guidebook. Version 2020. Available at https://www.acgme.org/globalassets/milestonesguidebook.pdf. Accessed 13 May 2023.
4. Tekian A, Hodges BD, Roberts TE, Schuwirth L, Norcini J. Assessing competencies using milestones along the way. Med Teach. 2015;37(4):399–402.
5. Englander R, Frank JR, Carraccio C, Sherbino J, Ross S, Snell L, on behalf of ICBME Collaborators. Toward a shared language for competency-based education. Med Teach. 2017;39(6):582–587.
6. Englander R, Burke AE, Guralnick S, Benson B, Hicks PJ, Ludwig S, et al. The Pediatric Milestones: a continuous quality improvement project is launched-now the hard work begins! Acad Pediatr. 2012;12(6):471–4.
7. Ten Cate O, Scheele F. Competency-based postgraduate training: can we bridge the gap between theory and clinical practice? Acad Med. 2007;82(6):542–7.
8. Ten Cate O. Nuts and bolts of entrustable professional activities. J Grad Med Educ. 2013;5(1):157–8.
9. Carraccio C, Englander R, Holmboe ES, Kogan J. Driving care quality: aligning trainee assessment and supervision through practical application of entrustable professional activities, competencies, and milestones. Acad Med. 2016;91(2):199–203.
10. Caverzagie KJ, Cooney TG, Hemmer PA, Berkowitz L. The development of entrustable professional activities for internal medicine residency training: a report from the Education Redesign Committee of the Alliance for Academic Internal Medicine. Acad Med. 2015;90:479–84.
11. Shaughnessy AF, Sparks J, Cohen-Osher M, Goodell K, Sawin G, Gravel J Jr. Entrustable professional activities in family medicine. J Grad Med Educ. 2013;5:112–8.
12. Ten Cate O, Schwartz A, Chen HC. Assessing trainees and making entrustment decisions: on the nature and use of entrustment-supervision scales. Acad Med. 2020;95:1662–99.
13. Crossley J, Johnson G, Booth J, Wade W. Good questions, good answers: construct alignment improves the performance of workplace-based assessment scales. Med Educ. 2011;45:560–9.
14. Schumacher D, West D, Schwartz A, Li ST, Millstein L, Griego E, et al. Longitudinal assessment of resident performance using Entrustable Professional Activities. JAMA Netw Open. 2020;3(1): e1919316.
15. Jones MD, Rosenberg A, Gilhooly J, Carraccio C. Perspective: Competencies, outcomes, and controversy—Linking professional activities to competencies to improve resident Education and practice. Acad Med. 2011;86(2):161–5.
16. Aylward M, Nixon J, Gladding S. An Entrustable Professional Activity (EPA) for handoffs as a model for EPA assessment development. Acad Med. 2014;89:1335–40.

17. Van der Vleuten CPM, Schuwirth LWT, Driessen EW, Dijkstra J, Tigelaar D, Baartman LKJ, van Tartwijk J. A model for programmatic assessment fit for purpose. Med Teach. 2012;34(3):205–14.
18. ABMS Insights. ABMS Member Boards Collaborate to Explore CBME. Available at: https://www.abms.org/news-events/abms-member-boards-collaborate-to-explore-cbme/. Accessed 31 May 2023.
19. American Board of Surgery. Training and Certification. Available at: https://www.absurgery.org/default.jsp?epahome. Accessed 31 May 2023.
20. Carraccio C, Englander R, Gilhooly J, Mink R, Hofkosh D, Barone M, et al. Building a framework of Entrustable Professional Activities, supported by competencies and milestones, to bridge the educational continuum. Acad Med. 2017;92(3):324–30.
21. Mink R, Schwartz A, Carraccio C, High P, Dammann C, McGann K, et. al. and the Steering Committee of the Subspecialty Pediatrics Investigator Network. Creating the subspecialty pediatrics investigator network (SPIN) J Pediatr. 2018;192:3–4,e2.
22. American Board of Pediatrics. Entrustable Professional Activities for subspecialties. https://www.abp.org/content/entrustable-professional-activities-subspecialties. Accessed 31 May 2023.
23. Mink R, Carraccio C, Herman B, Turner D, Myers A, Kesselheim J, et. al. and the Steering Committee of the Subspecialty Pediatrics Investigators Network (SPIN). Validity of level of supervision scales for assessing pediatric fellows on the common pediatric subspecialty Entrustable Professional Activities Acad Med. 2018;93(2):283–291.
24. Mink R, Herman B, Carraccio C, Aye T, Baffa J, Chess P, et. al. Subspecialty Pediatrics Investigator Network. Agreement of program directors with clinical competency committees for fellow entrustment. J Med Educ Curric Dev. 2020;7:2382120520936613.
25. Schwartz A, Young R, Hicks PJ. APPD LEARN. Medical education practice-based research networks: Facilitating collaborative research. Med Teach. 2014;38(1):64–74.
26. Mink R. Longitudinal evaluation of the required level of supervision for pediatric fellows. Available at https://www.appd.org/resources-programs/educational-resources/appd-learn/. Accessed 23 Aug 2023.
27. American Board of Pediatrics. Yearly growth in pediatric fellows by subspecialty by demographics and program characteristics. Available at: https://www.abp.org/content/yearly-growth-pediatric-fellows-subspecialty-demographics-and-program-characteristics. Accessed 31 May 2023.
28. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods. 1996;1(2):130–49.
29. Larrabee JG, Agrawal D, Trimm F, Ottolini M. Entrustable Professional Activities: Correlation of entrustment assessments of pediatric residents with concurrent subcompetency milestones ratings. J Grad Med Educ. 2020;12(1):66–73.
30. Turner DT, Schwartz A, Carraccio C, Herman B, Weiss P, Baffa J, et al. Continued supervision for the common pediatric subspecialty Entrustable Professional Activities may be needed following fellowship graduation. Acad Med. 2021;96(7S):S22–8.
31. Weiss P, Schwartz A, Carraccio C, Herman B, Mink R. Minimum supervision levels required by program directors for pediatric pulmonary fellow graduation. ATS Sch. 2021;2(3):360–369. https://doi.org/10.34197/ats-scholar.2021-0013OC.
32. Weiss P, Schwartz A, Carraccio C, Herman B, Turner D, Aye T, et al. Achieving entrustable professional activities during fellowship. Pediatrics. 2021;148(5): e2021050196.
33. Langhan M, Stafford D, Myers A, Herman B, Curran M, Czaja A, et al. Clinical competency committee perceptions of entrustable professional activities and their value in assessing fellows: a qualitative study of pediatric subspecialty program directors. Med Teach. 2023;45(6):650–7.
34. Liu L, Jiang Z, Qi X, Xie A, Wu H, Cheng H, et al. An update on current EPAs in graduate medical education: a scoping review. Med Educ Online. 2021;26(1):1981198.
35. Czaja A, Mink R, Turner D, Curran M, Herman B, Myers A, et. al. Facilitators and barriers to using entrustable professional activities in pediatric fellowships. Available at https://2022.pas-meeting.org/searchGlobal.asp?mode=Posters&SearchQuery=czaja. Accessed 31 May 2023.
36. Meier A, Gruessner A, Cooney R. Using the ACGME milestones for resident self-evaluation and faculty engagement. J Surg Educ. 2016;73:e150–7.

37. Schumacher D, Lewis K, Burke A, Smith ML, Schumacher J, Pitman M, et. al. The Pediatric Milestones: Initial evidence for their use as learning road maps for residents. Acad Pediatr. 2013(1);13:40–47.
38. Carraccio C, Burke A. Beyond competencies and milestones: adding meaning through context. J Grad Med Educ. 2010;2(3):419–422.
39. Li ST, Tancredi D, Schwartz A, Guillot A, Burke A, Trimm RF, et al. Competent for unsupervised practice: use of pediatric residency training milestones to assess readiness. Acad Med. 2017;92(3):385–93.
40. Yamazaki K, Holmboe E, Sangha S. Milestones PPV national report 2022. Accreditation Council for Graduate Medical Education. Available at https://www.acgme.org/globalassets/pdfs/milestones/acgmeppvreport2022.pdf. Accessed 31 May 2023.
41. Holmboe E, Yamazaki K, Nasca T, Hamstra S. Using longitudinal milestones data and learning analytics to facilitate the professional development of residents: early lessons from three specialties. Acad Med. 2020;95(1):97–103.
42. Carraccio C, Gusic M, Hicks MD. The Pediatrics Milestone Project. Acad Pediatr. 2014;14(2):S1–98.

Publisher's Note