BMC Medical Education

RESEARCH

Open Access

Entrustable professional activities for pediatric and subspecialties residency training in China

Yun Zhang[1,2†], Aijun Zhang[1†], Xiaoyu Jiao[1], Dongxiu Zhao[2], Yuankai Zhang[2], Tong Yue[2], Ke Yang[2] and Cuifen Zhao[1*]

Abstract

Background and objectives Outcome-based medical education is the latest focus in the past decade, and Entrustable Professional Activities(EPAs) have emerged as efficient vehicles to assess physicians. However, few studies have discussed the use of EPAs for residency training in pediatric medicine and its subspecialties. We conducted a pilot study to examine the feasibility of EPAs as a component of the clinical program of assessment in pediatric standardized residency training.

Methods We conducted a cross-sectional study for standardized residency training in different subspecialties within pediatric medicine at Qilu Hospital of Shandong University. Totally 65 residents and 35 directors joined in this study. An electronic EPA survey using 8 scales composed of 15 categories was distributed among residents and directors. Kruskal-Wallis test and Mann-Whitney U test were applied for comparing the self-assessments and director-assessments. Generalized estimated equation (GEE) was used to analyze the effect of postgraduate year(PGY), gender, and position on the EPA scores of director assessments.

Results A total of 401 director-assessment and 65 residents' self-assessment (response rate 100%) questionnaires were collected, both demonstrating rising trends in scores across PGYs. Significant differences were found between PGY1 and PGY2 ($p < 0.01$) and between PGY1 and PGY3 ($p < 0.01$), but not between PGY2 and PGY3 ($p > 0.01$). With an effect analysis of PGY, gender, and position on EPA scores performed, PGY had a significant effect on 13 out of 15 EPA scores, while gender affected only four EPA scores significantly, and position affected only three EPA categories. Meanwhile, some EPA categories revealed significant differences across various pediatric subspecialties ($p < 0.01$).

Conclusions The study findings suggest that EPA assessments is feasible among different PGYs in standardized Chinese residency training in pediatric medicine and its subspecialties. Postgraduate year had a significant impact on EPA scores, while gender and resident position also affected EPA scores to a certain extent. Improved stratified teaching programs are required for better subspecialty consistency.

Keywords Entrustable professional activities, Pediatric medicine, Pediatric subspecialties, Chinese standardized residency training, Director-assessment, Resident's self-assessment

†Yun Zhang and Aijun Zhang contributed equally to this work.

*Correspondence:
Cuifen Zhao
zhaocuifen@sdu.edu.cn
[1]Department of Pediatrics, Qilu Hospital of Shandong University, West Wenhua Road, Jinan 250012, China
[2]Medical Training Office, Qilu Hospital of Shandong University, Jinan, China

Article Summary

An EPAs survey was conducted to assess the clinical performance of residents and directors of Chinese pediatric standardized residency training in different subspecialties.

What's Known on This Subject

EPA assessments have a certain discriminating capability among different PGYs in Chinese standardized residency training in pediatric medicine and its subspecialties. Postgraduate year, gender and resident position affected EPA scores to a certain extent. Well feedback program is needed greatly.

What This Study Adds

Few studies have discussed the use of EPAs in pediatric residency training. We conducted a pilot study of the EPA for pediatric medicine and its subspecialties to evaluate the use of EPAs for residency training in pediatric medicine.

Introduction

Outcome-based medical education has been the focus of research in recent decades. A challenge faced by universities and other educational institutions when working within competency frameworks lies in the assessment of performance, particularly in specific workplace settings. Hence Entrustable Professional Activities(EPAs) have emerged as efficient vehicles for assessing physicians [1, 2, 3]. EPAs are defined as "units of professional practice, defined as tasks or responsibilities to be entrusted to unsupervised execution by a trainee once he or she has attained sufficient specific competence" [4]. The concept of "entrustable" refers to the ability of a physician to perform medical activities safely and effectively without supervisions [5]. EPAs are now widely used for the assessment of physicians in numerous training programs and curriculum development worldwide [6, 7, 8, 9, 10, 11]. Despite the seemingly universal affinity for EPAs, it is still not completely integrated into the assessment of resident progress and readiness for independent practice, especially in pediatric residency training.

The EPA concept has been rapidly adopted in different specialties in postgraduate medical education, including, but not limited to, psychiatry, radiology, gastroenterology, pathology, pediatrics and palliative medicine [12, 13, 14, 15, 16, 17, 18, 19, 20]. In recent years, there has been a growing body of research on the application of EPA in pediatric professional training and demonstrated that EPA assessment was both valid and reliable [21]. However, the vast majority of studies focused on fellowship rather than residency [22, 23, 24, 25, 26, 27]. Furthermore, there is still little research of the EPA for pediatric residency training in China and the study lacked comparisons between subspecialties [28]. In China, Peking University First Hospital evaluated a formative assessment system based on EPAs in pediatric residency training, proposing an EPA system to assess postgraduate medical education (PGME) that was made up of 15 EPA categories in eight scales(translated as Tables 1 and 2). This highlights the complementary advantage of EPAs that could be integrated with the ongoing competency-based medical education(CBME) formative assessment program. CBME is an outcomes-based education approach that involves identifying the abilities needed by the physician and designing the curriculum to both support and assess the attainment of these predetermined competencies. It includes mini-clinical-evaluation exercises (Mini-CEX), direct observation of procedural skills (DOPS), subjective-objective-assessment-plan (SOAP), and 360-degree assessment. Hence, we began to push forward

Table 1 EPAs categories

Number	Category
1	Admit a patient
2	Select and interpret auxiliary examinations
3	Diagnose and make the differential diagnosis
4	Make therapeutic decision
5	Compose medical documents
6	Report a case
7	Recognize and manage general clinical conditions
8	Recognize and manage emergent and critical conditions
9	Transfer and hand over a patient
10	Perform informed consent
11	Perform basic operation
12	Perform health education
13	Inform bad news
14	Perform clinical education
15	Manage public health events

Table 2 Eight entrustable levels of each EPA

Scale	Details
1	Can not perform certain professional activities as a resident under the direct supervision of a superior physician
2	Perform certain professional activities with a superior physician together
3	Perform certain professional activities under the supervision and guidance of a superior physician
4	Perform certain professional activities without the presence of a superior physician; when help is needed, need the presence of a superior physician to recheck all performances.
5	Perform certain professional activities without the presence of a superior physician; when help is needed, need the presence of a superior physician to recheck important performances.
6	Perform certain professional activities without the presence of the superior physician; when help is needed, need the guidance and recheck of superior physician over the phone.
7	Perform certain professional activities without the need for supervision and guidance from a superior physician.
8	Can provide supervision and guidance for others in certain professional activities.

an EPA assessment program in the Pediatric Department at Qilu Hospital in Shandong University based on the CBME system for standardized residency training at the Peking University First Hospital.

Our previous study indicated that EPA assessments have a certain discriminating capability among different PGYs in Chinese standardized residency training in pediatric intensive care medicine, and postgraduate year, gender, and resident position affected EPA scores to a certain extent [29]. A research from the United States also indicated that EPA scores in pediatric residents were correlated with PGY [30]. To further explore the effectiveness of EPAs and deficiencies in residency training in the entire pediatric department, we performed a pilot study of 65 residents within the Chinese standardized residency training program in eight subspecialties of the Pediatric Department at Qilu Hospital of Shandong University and solicited both resident self-assessment and director-assessment of this training model.

Based on previous research and the different spectrum of diseases between Pediatric Department subspecialties, we hypothesized that: 1)EPA assessments can be effectively applied to various PGYs in standardized Chinese residency training in pediatric medicine and its subspecialties. 2)EPA scores were significantly impacted by postgraduate year, while gender and resident position also affected EPA scores to a certain extent. 3)EPA scores showed differences across different subspecialties in the same PGYs.

Materials and methods

Setting

Like many other Chinese standardized training residency programs, Qilu Hospital of Shandong University has a CBME evaluation course that spans residents' training after graduation from medical school and includes Mini-CEX, DOPS, SOAP, and 360-degree assessment. According to the national guidelines for standardized residency training, every pediatric resident is supposed to rotate through all pediatric subspecialties(Pediatric Hematology and Oncology, Urology, Neurology, Respiratory, Neonatology, Cardiology, Gastroenterology, Infectious Diseases, Intensive Care Medicine, Endocrinology, Genetic Metabolism Diseases, Rheumatology, Immunology and Child Healthcare) within a 3-year training phase. We have certain subspecialties united into one department because of the structure of departmental administrative divisions. The departmental rotation examination is administered at the end of each rotation phase and is composed of all of the above-mentioned skill tests and formative assessments. The directors in charge of each subspecialty were pediatric physicians well trained by national or provincial director courses for Chinese standardized training residency programs who obtained qualification certifications from the Chinese Health Commission or Shandong Provincial Health Commission.

Sample

This study enrolled 65 residents who were trained in a standardized residency training program from August 2019 to July 2022 at the Qilu Hospital of Shandong University. Meanwhile, 35 directors in charge of each pediatric department over the same period were recruited for this study. All enrolled residents were categorized as postgraduate year 1 (PGY1) to PGY3, according to their seniority. Resident positions were categorized into four types(professional master, entrusted training residents, permanent staff at our hospital, and social training residents) according to their affiliations. Professional masters referred to graduate students who were concurrently enrolled in standardized residency training while pursuing a master's degree in clinical medicine. Entrusted training residents were those who were commissioned by the employing organization, so they generally possessed clinical experience. Permanent staff referred to new employees of our hospital who were undergoing standardized residency training. They had the highest educational background. Social training residents were those who were recruited through social recruitment into our training base. This study was approved by the Qilu Hospital of Shandong University's Institutional Review Board. This study has obtained the consent of all the participants and passed the review of the Qilu Hospital of Shandong University's ethics committee. Clinical trial number: not applicable.

Procedure

This cross-sectional study was conducted at the end of July, 2022. EPA resident self-assessments and director-assessments were used at the time of the research to evaluate resident performance and competency from both points of view in this moment. An electronic questionnaire composed of EPAs with 15 categories on eight scales was administered to solicit both resident self-assessment and director-assessment, in addition to the ongoing evaluation program (Mini-CEX, DOPS, SOAP, and 360-degree assessment). The director-assessment of each resident was performed by several directors, whereas the self-assessment of each resident was performed by the resident. Each questionnaire included general information (director name, resident name, resident gender, seniority, and position, such as professional masters, entrusted training residents from junior hospitals, residents of permanent staff at Qilu Hospital of Shandong University, and social training residents) and EPA evaluation. The 15 categories of EPA evaluation were established using the guidelines of Peking University First Hospital (translated as Table 1). Based on

previous literature, each EPA was set using eight scales (translated as in Table 2). All EPA assessments were performed until the participating residents or directors were well informed about the details of the questionnaire. All questionnaires were electronically administered using mobile software. Multiple reminders and phone follow-ups by data collection staff were set up to ensure that all required responses were collected in time. Each enrolled questionnaire indicated that all included questions were completed.

Statistical analysis

All questionnaires were administered using the Wenjuanwang APP 2.7.0(Zhongyan Network Technology Co., Ltd., Shanghai, China). Data were collected using Excel (Microsoft, Redwood, WA, USA), and statistical analysis and figure creation were performed using SPSS 23.0.0 (IBM, Armonk, NY, USA). Comparisons of self-assessments and director-assessments for every EPA across different PGYs were statistically analyzed using the Kruskal-Wallis test. Statistical significance was defined as a two-sided p-value of <0.05. Comparisons between self-assessments and director-assessments for every EPA between every two PGY levels were statistically analyzed using the Mann-Whitney U test, with significance defined as a corrected p-value of 0.017 using the Bonferroni correction three times the Mann-Whitney U test for the same EPA. The effect of PGY, gender, and position on the EPA scores of director assessments was analyzed using the generalized estimated equation (GEE), with $p<0.05$, considered statistically significant.

Results

General information

This study included 65 residents (see Table 3) and 35 directors. The results included 65 resident self-assessment questionnaires and 401 director-assessment questionnaires, with a 100% response rate. The characteristics of the enrolled participants are shown in Table 3. Of the 65 residents, 53 (81.5%) were female. PGY1 to PGY3 respectively accounted for 36.9%(24/65), 40.0%(26/65) and 23.1%(15/65). A line graph was created to show the trends in director and self-assessment EPA scores over progressive PGY levels (Fig. 1). A rising trend in director-assessment scores across all EPAs by PGY increase was noted, especially between PGY1 vs. PGY2 and PGY1 vs. PGY3. Whereas self-assessment scores showed a non-distinctive trend across different PGYs. The categories of the PGYs, genders, positions, and directors' subspecialties in the director-assessment questionnaires are summarized in Table 4.

Table 3 Characteristics of residents

Characteristics	PGY1	PGY2	PGY3	*p*-value
Number of residents, n(%)	24(36.9%)	26(40.0%)	15(23.1%)	-
Male, n(%)	5(20.8%)	5(19.2%)	2(13.3%)	0.83
Number of director-assessments, mean ± SD	4.3 ± 2.4[a]	6.4 ± 2.6[b]	8.7 ± 1.3[c]	***0.00***[*]

[*] Kruskal-Wallis test; [a]Mann-Whitney U test revealed a significant difference ($p=0.011$, $p<0.17$) between PGY1 and PGY2; [b]Mann-Whitney U test revealed a significant difference ($p=0.000$, $p<0.17$) between PGY2 and PGY3; [c]Mann-Whitney U test revealed a significant difference ($p=0.000$, $p<0.17$) between PGY1 and PGY3

Comparison of director-assessment scores across different PGYs, genders and positions

The Kruskal-Wallis test revealed significant differences between the EPA director-assessment scores across different PGYs (Table 5). The higher the PGY year in which the residents were in, the higher their scores in each EPA category. When univariate PGY years were compared, there were significant differences between PGY1 and PGY2, and between PGY1 and PGY3 ($p<0.017$), whereas there were no obvious differences between PGY2 and PGY3 in any EPA category. As to comparison between genders, there were no significant differences between genders except for EPA13(Inform bad news) ($p=0.024$), EPA15(Manage public health events)($p=0.042$)(Fig. 2) by Mann-Whitney U test, with male residents scoring higher than those of the female. When comparing the director-assessment scores across different positions, there were significant differences in all EPAs(EPA1: $p=0.09$, EPA3: $p=0.01$, EPA4:$p=0.020$, EPA5: $p=0.007$, EPA6: $p=0.009$, EPA7: $p=0.002$, EPA8: $p=0.002$, EPA9: $p=0.006$, EPA10: $p=0.000$, EPA11: $p=0.002$, EPA12: $p=0.000$, EPA13: $p=0.000$, EPA14: $p=0.012$, EPA15: $p=0.026$, $p<0.05$) except for EPA2(select and interpret auxiliary examinations) ($p=0.050$) using the Kruskal-Wallis test.

Effect analysis of PGY, gender, and position on director-assessment scores

To analyze the effect of residents' PGY, gender, and position on EPA scores, a generalized estimated equation (GEE) model analysis was performed (Table 6). PGY had a significant effect on all EPA scores ($p<0.05$), except for EPA5 ($p=0.066$,>0.05) and EPA7 ($p=0.128$,>0.05), whereas the resident's position significantly affected only EPA10 ($p=0.003$,<0.05), EPA13 ($p=0.000$,<0.05), and EPA14 ($p=0.001$,<0.05). Residents' gender was only correlated with EPA8 ($p=0.014$,<0.05), EPA10 ($p=0.023$,<0.05), EPA13 ($p=0.036$,<0.05), and EPA14 ($p=0.010$,<0.05)(Table 6).

The scores of all 13 significantly correlated EPA categories rose as PGY grew except for EPA2 (Select and interpret auxiliary examinations, set PGY1 as zero; PGY2:B = 0.711, $p=0.040$, <0.05, PGY3:B = 0.552, $p=0.524$, >0.05), EPA6 (Report a case, PGY2:B = 1.079 $p=0.029$, <0.05, PGY3:B = 0.813, $p=0.000$, <0.05),

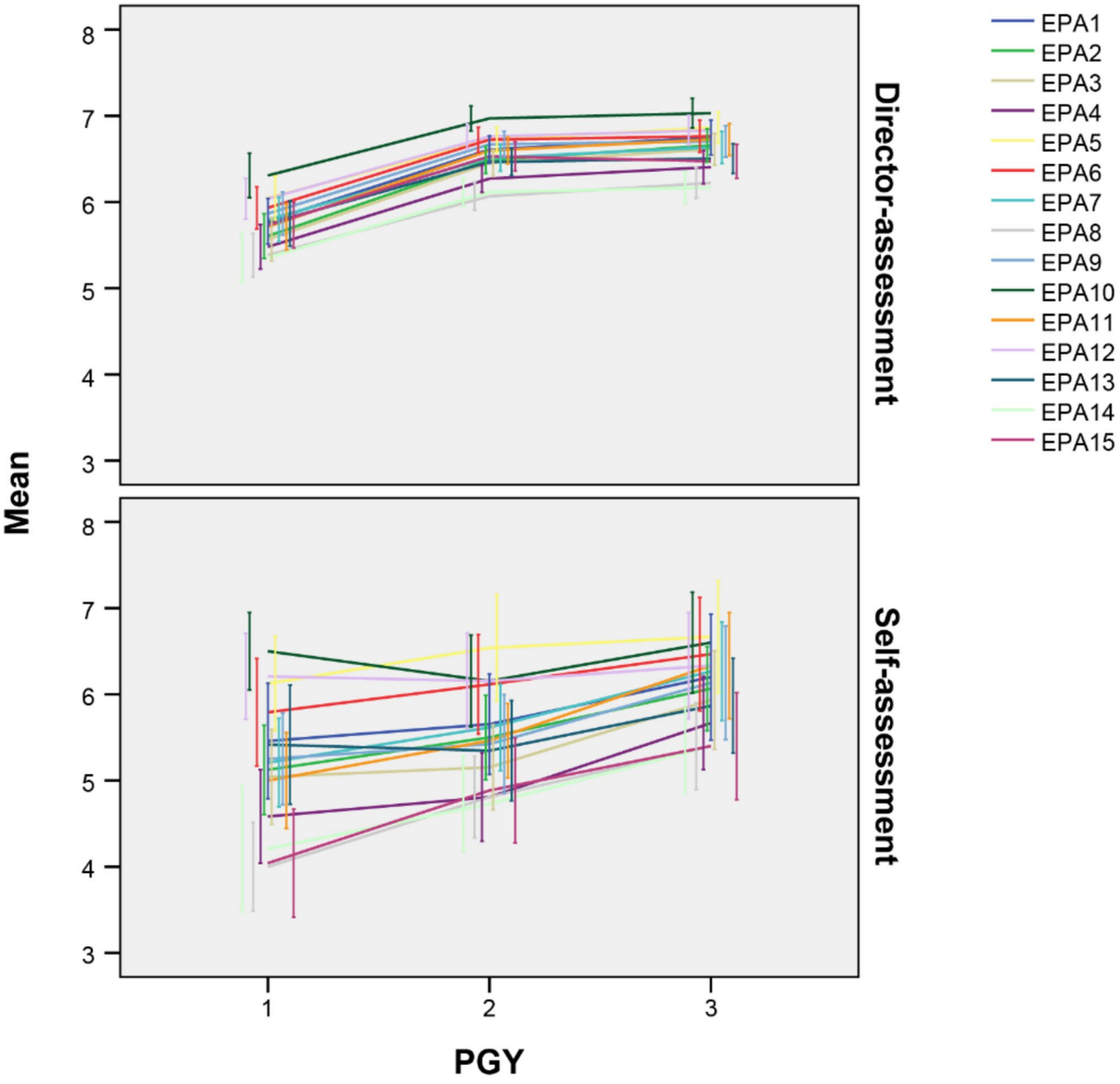

Fig. 1 Line graph of scores of director- and self-assessment in each EPA*
*Each point represents the mean score of a certain subgroup, with bars representing the 95% CI of the mean of each subgroup

EPA14(Perform clinical education, PGY2:B = 0.900, $p = 0.000$, < 0.05, PGY3:B = 0.693, $p = 0.000$, < 0.05) and EPA15(Manage public health events, PGY2:B = 0.913, $p = 0.000$, < 0.05, PGY3:B = 0.850, $p = 0.000$, < 0.05), with higher mean scores for PGY2s than PGY3s and the lowest mean score at PGY1.

Regarding the scores of the four EPAs with significant gender gap, EPA8 (recognize and manage emergent and critical conditions, set Female as zero; Male: B = 0.271, $p = 0.014$, < 0.05), EPA10 (Perform informed consent, Male: B = 0.343, $p = 0.023$, < 0.05), EPA13 (inform bad news, Male: B = 0.337, $p = 0.036$, < 0.05), and EPA14 (Perform clinical education, Male: B = 1.040, $p = 0.010$, < 0.05) showed that male residents scored significantly higher than female residents.

Meanwhile, for the significant resident position effect on the three EPAs, permanent staff scored highest in EPA10 (Perform informed consent, set professional master as zero; Permanent staff B = 0.500, $p = 0.042$, < 0.05; entrusted training residents B = 0.426, $p = 0.001$, < 0.05, social training residents B = 0.113, $p = 0.428$, > 0.05) and EPA13 (Inform bad news, set professional master as zero; Permanent staff B = 0.411, $p = 0.000$, < 0.05, Entrusted training residents B = 0.321, $p = 0.010$, < 0.05, Social

Table 4 Categorical variable information of director-assessment questionnaires

Factor		*N*	Percent
PGY	PGY1	104	25.9%
	PGY2	166	41.4%
	PGY3	131	31.7%
Gender	Female	323	80.5%
	Male	78	19.5%
Position	Professional master	204	50.9%
	Entrusted training residents	159	39.7%
	Permanent staff	18	4.5%
	Social training residents	20	5.0%
Subspecialty	Pediatric Neurology	22	5.5%
	Pediatric Hematology	44	11.0%
	Pediatric Respiratory	29	7.2%
	Pediatric Gastroenterology	24	6.0%
	Pediatric Cardiology	28	7.0%
	Pediatric Urology	26	6.5%
	Pediatric Intensive Care Unit	113	28.2%
	Neonatal Intensive Care Unit	88	21.9%
	Child Healthcare	27	6.7%

Table 5 Scores of director-assessment in different PGYs

EPAs	PGY1	PGY2	PGY3	Chi-square*	*P*-value
EPA1	5.8 ± 1.4**	6.6 ± 1.0	6.7 ± 1.2***	37.037	***0.000***
EPA2	5.6 ± 1.3**	6.5 ± 1.0	6.7 ± 1.1***	44.055	***0.000***
EPA3	5.6 ± 1.3**	6.5 ± 1.0	6.6 ± 1.1***	46.339	***0.000***
EPA4	5.5 ± 1.3**	6.3 ± 1.0	6.4 ± 1.1***	36.348	***0.000***
EPA5	6.0 ± 1.3**	6.7 ± 1.0	6.9 ± 1.0***	29.547	***0.000***
EPA6	5.9 ± 1.2**	6.7 ± 1.0	6.8 ± 1.1***	34.835	***0.000***
EPA7	5.8 ± 1.3**	6.5 ± 1.0	6.6 ± 1.1***	28.940	***0.000***
EPA8	5.4 ± 1.3**	6.1 ± 1.0	6.2 ± 1.0***	30.324	***0.000***
EPA9	5.9 ± 1.3**	6.7 ± 1.0	6.7 ± 1.1***	33.957	***0.000***
EPA10	6.3 ± 1.3**	7.0 ± 1.0	7.0 ± 1.0***	22.939	***0.000***
EPA11	5.7 ± 1.4**	6.6 ± 1.0	6.7 ± 1.1***	40.616	***0.000***
EPA12	6.0 ± 1.2**	6.8 ± 1.0	6.8 ± 1.0***	34.307	***0.000***
EPA13	5.8 ± 1.3**	6.5 ± 1.0	6.5 ± 1.0***	26.417	***0.000***
EPA14	5.4 ± 1.5**	6.1 ± 1.1	6.2 ± 1.1***	23.002	***0.000***
EPA15	5.7 ± 1.4**	6.5 ± 1.1	6.5 ± 1.1***	27.237	***0.000***

*Kruskal Wallis test; **Mann-Whitney U test revealed a significant difference (all P-values = 0.000)between PGY1 and PGY2; ***Mann-Whitney U test revealed a significant difference (all P-values = 0.000) between PGY1 and PGY3

training residents B=-0.030, $p = 0.903$, > 0.05). Entrusted training residents obtained the highest scores in EPA14 (Perform clinical education, set professional master as zero; entrusted training residents B = 0.622, $p = 0.004$, < 0.05; Permanent staff B = 0.576, $p = 0.009$, < 0.05; social training residents B=-0.112, $p = 0.443$, > 0.05).

Comparison of director-assessment scores across different subspecialties

The comparison of director-assessment scores across different subspecialties showed significant differences in all EPAs(all $p = 0.000$,<0.05). For each PGY, the scores from different director subspecialties showed significant differences within PGY1 in all EPAs except for EPA11(Perform basic operation, $p = 0.074$, > 0.05), within PGY2 in all EPAs except for EPA 8(Recognize and manage emergent and critical conditions, $p = 0.052$, > 0.05), EPA10(Perform informed consent, $p = 0.050$, > 0.05), and EPA12(Perform health education, $p = 0.072$, > 0.05), within PGY3 in EPA1(Admit a patient, $p = 0.166$, > 0.05), EPA2(Select and interpret auxiliary examinations, $p = 0.050$, > 0.05), and EPA10(Perform informed consent, $p = 0.053$, > 0.05) (Fig. 3).

Comparison of director-assessment scores from pediatric intensive care medicine and neonatology

Among all the subspecialties of directors enrolled in this research, the number of recruited questionnaires from Pediatric Intensive Care Medicine and Neonatology ranked the top two among all subspecialties. In the comparison of the scores from these two intensive care medicines, within all PGYs there were no significant differences except for EPA9(transfer and hand over a patient, $p = 0.033$, < 0.05), EPA11(Perform basic operation, $p = 0.044$, < 0.05), and EPA15(manage public health events, $p = 0.037$, < 0.05) with scores from Pediatric Intensive Care Medicine higher than those from neonatology. However, within each single PGY, there were no obvious differences between the two subspecialties($p > 0.05$) (Fig. 4).

Comparison of self-assessment scales across different PGYs

For the self-assessment EPA scores of all enrolled residents, there were significant differences only within EPA2 (select and interpret auxiliary examinations), EPA4 (make therapeutic decisions), EPA7 (recognize and manage general clinical conditions), EPA8 (recognize and manage emergent and critical conditions), EPA11 (perform basic operations), and EPA15 (manage public health events) across the different PGYs, with higher levels of PGY residents scoring better (Table 7). As for the comparisons between the two PGYs, there were no significant differences in any EPA scores between PGY1 and PGY2 and PGY2 and PGY3(all $p > 0.017$). However, significant differences in EPA7 (recognition and management of general clinical conditions, $p = 0.010$,<0.017), EPA8 (recognizing and managing emergent and critical conditions, $p = 0.001$,<0.017), EPA11 (Perform basic operation, $p = 0.003$,<0.017), and EPA15 (managing public health events, $p = 0.015$,<0.017) were observed between PGY1 and PGY3 ($p < 0.017$)(Fig. 1).

Comparison of EPAs scores of self-assessment between genders

There were no significant differences in EPA between the self-assessment scores of male and female residents.

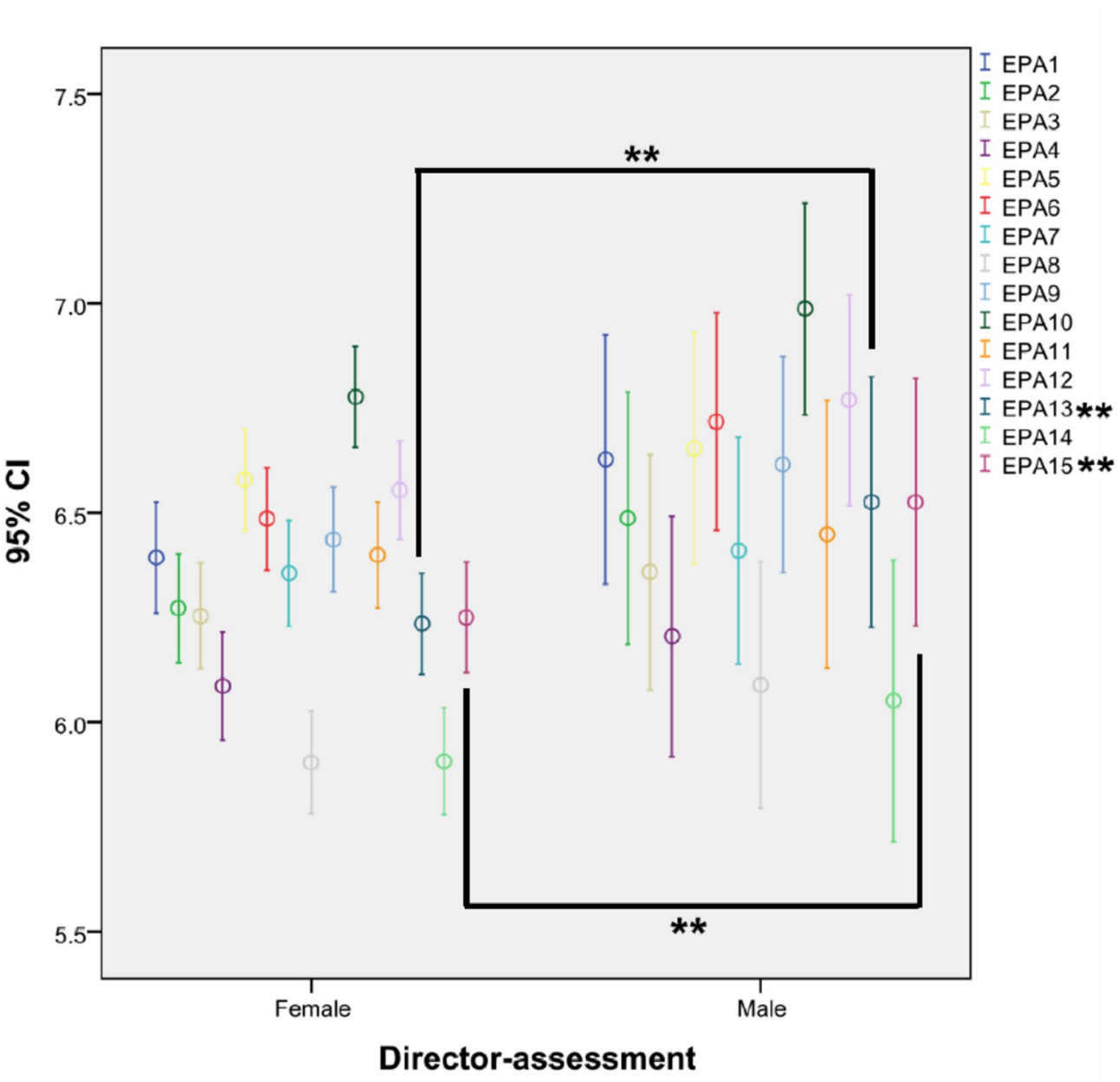

Fig. 2 Error bar chart of director-assessment between genders
*The edges of each bar represent the 95% CI of the scores in the subgroups. **EPA13: $p = 0.024$, $p < 0.05$; EPA15: $p = 0.042$, $p < 0.05$.

Comparisons between director and self-assessment scores across EPAs within the same PGY

The comparisons between director and self-assessment scores across EPAs are presented in Fig. 5. The director and self-assessment scores of PGY1s were mostly consistent, except for EPA4 (make therapeutic decision, $p = 0.007$, $p < 0.05$), EPA8 (recognize and manage emergent and critical conditions, $p = 0.000$, $p < 0.05$), EPA11 (Perform basic operation form, $p = 0.015$, $p < 0.05$), EPA14 (Perform clinical education, $p = 0.002$, $p < 0.05$), and EPA15 (manage public health events, $p = 0.000$, $p < 0.05$), which directors awarded higher scores. There were significant differences between the self-assessment and director-assessment scores for every EPA for PGY2s, except for EPA5 (composed medical documents, $p = 0.762$, $p > 0.05$). For PGY3s, the director and self-assessment scores of PGY1s were consistent, except for EPA2 (select and interpret auxiliary examinations, $p = 0.031$, $p < 0.05$), EPA3 (diagnose and make the differential diagnosis, $p = 0.022$, $p < 0.05$), EPA4 (make therapeutic decision,

Table 6 Generalized estimated equation analysis of director-assessment questionnaires

EPAs	Factor	Tests Of Model Effects		Parameter	B	95% Wald Confidence Interval			Hypothesis Test
		Ward Chi-Square	P-value			Lower	Upper	Wald Chi-Square	P-value
EPA1	**PGY**	14.066	***0.001***	**PGY1**	0[a]				
				PGY2	0.920	0.425	1.414	5.424	***0.020***
				PGY3	1.195	0.189	2.202	13.294	***0.000***
	Gender	3.079	0.079	**Female**	0[a]				
				Male	0.329	-0.038	0.696	3.079	0.079
	Position	5.341	0.148	**Professional master**	0[a]				
				Entrusted training residents	0.182	-0.397	0.761	0.761	0.378
				Social training residents	-0.304	-0.966	0.359	0.806	0.369
				Permanent staff	-0.086	-0.833	0.660	0.051	0.821
EPA2	**PGY**	9.183	***0.010***	**PGY1**	0[a]				
				PGY2	0.711	0.033	1.390	4.221	***0.040***
				PGY3	0.552	-1.148	2.253	0.405	0.524
	Gender	1.825	0.177	**Female**	0[a]				
				Male	0.421	-0.190	1.033	0.992	0.177
	Position	2.279	0.517	**Professional master**	0[a]				
				Entrusted training residents	0.473	-0.458	1.405	0.992	0.319
				Social training residents	0.058	-0.910	1.026	0.014	0.907
				Permanent staff	0.578	-0.823	1.980	0.654	0.419
EPA3	**PGY**	19.936	***0.000***	**PGY1**	0[a]				
				PGY2	0.732	0.392	1.072	17.811	***0.000***
				PGY3	0.913	0.450	1.376	14.964	***0.000***
	Gender	0.001	0.979	**Male**	0[a]				
				Female	-0.005	-0.397	0.387	0.001	0.979
	Position	6.792	0.079	**Professional master**	0[a]				
				Entrusted training residents	0.246	-0.092	0.584	2.038	0.153
				Social training residents	-0.068	-0.535	0.398	0.082	0.775
				Permanent staff	0.467	0.051	0.884	4.848	***0.028***
EPA4	**PGY**	14.689	***0.001***	**PGY1**	0[a]				
				PGY2	0.651	0.280	1.022	11.841	***0.001***
				PGY3	0.776	0.350	1.202	12.767	***0.000***
	Gender	0.680	0.410	**Female**	0[a]				
				Male	-0.211	-0.586	0.164	1.219	0.270
	Position	3.271	0.352	**Professional master**	0[a]				
				Entrusted training residents	0.277	-0.054	0.608	2.683	0.101
				Social training residents	0.238	-0.218	0.694	1.045	0.307
				Permanent staff	0.341	-0.133	0.814	1.989	0.158
EPA5	**PGY**	5.421	0.066	**PGY1**	0[a]				
				PGY2	0.314	-0.602	1.230	0.452	0.501
				PGY3	0.756	0.117	1.396	5.370	***0.020***
	Gender	0.237	0.626	**Female**	0[a]				
				Male	-0.571	-2.870	1.728	0.237	0.626
	Position	0.187	0.980	**Professional master**	0[a]				
				Entrusted training residents	-0.001	-1.099	1.097	0.000	0.999
				Social training residents	-0.008	-0.478	0.461	0.461	***0.001***
				Permanent staff	0.154	-0.661	0.968	0.968	0.137

Table 6 (continued)

EPAs	Factor	Tests Of Model Effects		Parameter	B	95% Wald Confidence Interval			Hypothesis Test
		Ward Chi-Square	P-value			Lower	Upper	Wald Chi-Square	P-value
EPA6	**PGY**	20.678	***0.000***	**PGY1**	0[a]				
				PGY2	1.079	0.111	2.046	4.776	***0.029***
				PGY3	0.813	0.450	1.177	19.224	***0.000***
	Gender	0.803	0.370	**Female**	0[a]				
				Male	1.203	-1.429	3.834	0.803	0.370
	Position	1.239	0.744	**Professional master**	0[a]				
				Entrusted training residents	0.662	-0.580	1.903	1.092	0.296
				Social training residents	0.047	-0.415	0.508	0.040	0.842
				Permanent staff	0.312	-0.367	0.991	0.809	0.368
EPA7	**PGY**	4.113	0.128	**PGY1**	0[a]				
				PGY2	0.596	0.015	1.177	4.049	***0.044***
				PGY3	0.647	-0.224	1.519	2.122	0.145
	Gender	0.092	0.761	**Female**	0[a]				
				Male	0.171	-0.934	1.277	1.858	0.173
	Position	1.626	0.654	**Professional master**	0[a]				
				Entrusted training residents	0.409	-0.296	1.113	1.292	0.256
				Social training residents	0.037	-0.501	0.574	0.018	0.894
				Permanent staff	0.269	-0.384	0.922	0.653	0.419
EPA8	**PGY**	34.619	***0.000***	**PGY1**	0[a]				
				PGY2	0.668	0.655	1.313	18.042	***0.000***
				PGY3	0.984	0.360	0.977	34.378	***0.000***
	Gender	5.978	***0.014***	**Female**	0[a]				
				Male	0.271	0.054	0.488	5.978	***0.014***
	Position	6.891	0.075	**Professional master**	0[a]				
				Entrusted training residents	0.286	0.027	0.545	4.676	***0.031***
				Social training residents	0.234	-0.762	0.035	0.704	0.401
				Permanent staff	0.186	-0.330	0.701	0.499	0.480
EPA9	**PGY**	41.883	***0.000***	**PGY1**	0[a]				
				PGY2	0.846	0.468	1.224	19.211	***0.000***
				PGY3	1.014	0.701	1.327	40.338	***0.000***
	Gender	0.218	0.641	**Female**	0[a]				
				Male	0.185	-0.591	0.960	0.218	0.641
	Position	1.758	0.624	**Professional master**	0[a]				
				Entrusted training residents	0.144	-0.276	0.564	0.564	0.451
				Social training residents	-0.204	-0.717	0.309	0.607	0.436
				Permanent staff	-0.069	-0.377	0.239	0.191	0.662
EPA10	**PGY**	16.163	***0.000***	**PGY1**	0[a]				
				PGY2	0.578	0.242	0.915	11.348	***0.001***
				PGY3	0.689	0.352	1.027	16.039	***0.000***
	Gender	5.163	***0.023***	**Female**	0[a]				
				Male	0.343	0.047	0.640	5.163	***0.023***
	Position	13.772	***0.003***	**Professional master**	0[a]				
				Entrusted training residents	0.426	0.176	0.676	11.128	***0.001***
				Social training residents	0.113	-0.167	0.393	0.627	0.428
				Permanent staff	0.500	0.017	0.987	4.123	***0.042***

Table 6 (continued)

EPAs	Factor	Tests Of Model Effects		Parameter	B	95% Wald Confidence Interval			Hypothesis Test
		Ward Chi-Square	P-value			Lower	Upper	Wald Chi-Square	P-value
EPA11	**PGY**	32.099	***0.000***	**PGY1**	0[a]				
				PGY2	0.813	0.432	1.193	17.528	***0.000***
				PGY3	1.043	0.675	1.410	30.923	***0.000***
	Gender	0.036	0.849	**Female**	0[a]				
				Male	0.065	-0.608	0.739	0.036	0.849
	Position	2.255	0.521	**Professional master**	0[a]				
				Entrusted training residents	0.281	-0.112	0.675	1.967	0.161
				Social training residents	-0.064	-0.605	0.477	0.054	0.816
				Permanent staff	0.157	-0.180	0.494	0.936	0.361
EPA12	**PGY**	10.033	***0.007***	**PGY1**	0[a]				
				PGY2	0.355	-0.278	0.988	1.207	0.272
				PGY3	0.645	0.246	1.045	10.032	***0.002***
	Gender	0.190	0.663	**Female**	0[a]				
				Male	-0.349	-1.915	1.218	0.190	0.663
	Position	1.346	0.718	**Professional master**	0[a]				
				Entrusted training residents	0.103	-0.656	0.863	0.071	0.790
				Social training residents	0.153	-0.180	0.487	0.811	0.368
				Permanent staff	0.027	-0.288	0.343	0.028	0.866
EPA13	**PGY**	16.484	***0.005***	**PGY1**	0[a]				
				PGY2	0.612	0.277	0.947	12.814	***0.000***
				PGY3	0.718	0.366	1.070	15.960	***0.000***
	Gender	4.409	***0.036***	**Female**	0[a]				
				Male	0.337	0.022	0.652	4.409	***0.036***
	Position	18.602	***0.000***	**Professional master**	0[a]				
				Entrusted training residents	0.321	0.073	0.569	6.418	***0.010***
				Social training residents	-0.030	-0.517	0.457	0.015	0.903
				Permanent staff	0.411	0.209	0.612	15.964	***0.000***
EPA14	**PGY**	20.351	***0.000***	**PGY1**	0[a]				
				PGY2	0.900	0.470	1.331	16.797	***0.000***
				PGY3	0.693	0.340	1.045	14.857	***0.000***
	Gender	6.566	***0.010***	**Female**	0[a]				
				Male	1.040	0.245	1.836	6.566	***0.010***
	Position	16.719	***0.001***	**Professional master**	0[a]				
				Entrusted training residents	0.622	0.203	1.042	8.461	***0.004***
				Social training residents	-0.112	-0.397	0.173	0.590	0.443
				Permanent staff	0.576	0.143	1.009	6.786	***0.009***
EPA15	**PGY**	26.290	***0.000***	**PGY1**	0[a]				
				PGY2	0.913	0.514	1.312	20.101	***0.000***
				PGY3	0.850	0.516	1.184	24.871	***0.000***
	Gender	3.170	0.075	**Female**	0[a]				
				Male	0.567	-0.057	1.192	3.170	0.075
	Position	4.930	0.177	**Professional master**	0[a]				
				Entrusted training residents	0.315	-0.047	0.677	2.910	0.088
				Social training residents	-0.208	-0.598	0.182	1.091	0.296
				Permanent staff	0.095	-0.234	0.425	0.322	0.570

[a]Set to zero because this parameter is redundant

$p = 0.012$, $p < 0.05$), EPA8 (recognize and manage emergent and critical conditions, $p = 0.004$, $p < 0.05$), EPA13 (form bad news, $p = 0.026$, $p < 0.05$), EPA14 (Perform clinical education, $p = 0.011$, $p < 0.05$), and EPA15 (manage public health events, $p = 0.001$, $p < 0.05$), in which directors awarded higher scores(Fig. 5).

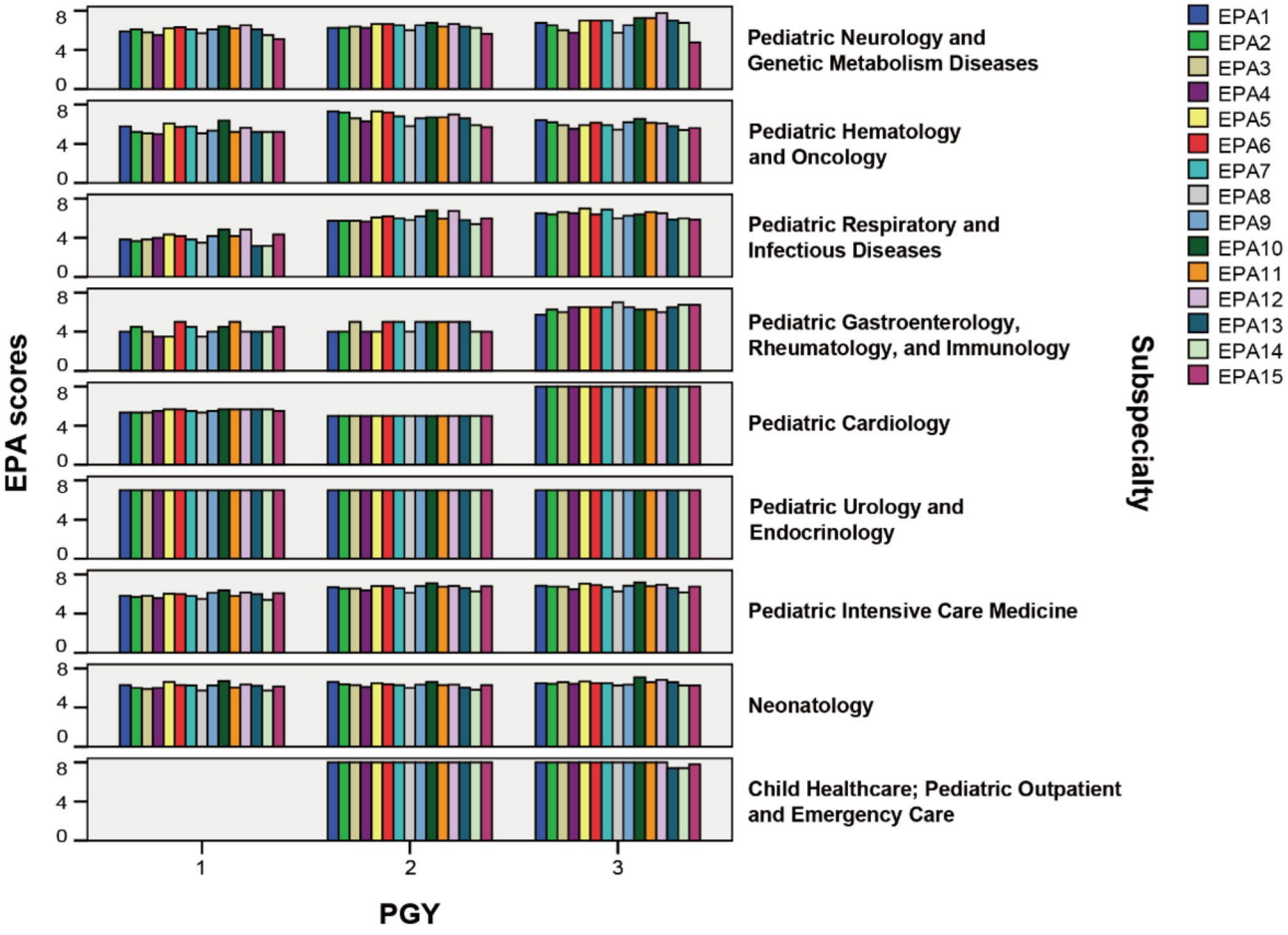

Fig. 3 Boxplot of director-assessment scores of EPAs in different subspecialties across PGYs

Discussion

Entrustable professional activities(EPAs) are designed to be real-life activities, and as such, can be understood and applied more easily than prior concepts within competency-based medical education(CBME), such as milestones [31]. Each EPA combines the knowledge, skills, and attitudes necessary to perform a medical task, and acts as a means of bridging the gap between these theoretical competencies and the assessment of competence. Moreover, it is an important tool for evaluating the completion of tasks and responsibilities unsupervised. EPAs provide a framework to make explicit judgments of trainee ability, which is important at all stages of medical education [32]. A review article indicated that in recent years, there has been a notable shift from descriptions of EPA development processes toward aspects beyond development, such as implementation, feasibility, acceptance/perception, and assessment [33]. Notably, there are few studies on EPAs in pediatric postgraduate education, let al.one the related subspecialties of pediatrics. Furthermore, studies on EPAs are mostly conducted in the United States, Australia, Germany and Canada [34, 35, 36, 37, 38], rather than Asia. Our study aimed to supplement the implementation and feasibility of EPAs and deficiencies in standardized Chinese residency training in pediatric medicine and its subspecialties.

This cross-sectional study evaluated pediatric residents trained in Chinese standardized residency training at the Qilu Hospital of Shandong University in July 2022. The sample size of this study was large, including 65 residents and 35 directors, and there was satisfying consistency among questionnaires. We focused on the implementation and feasibility of entrustable professional activities within multiple pediatric subspecialties and within the entire pediatric specialty. Similar to our previous study on EPAs in pediatric intensive care medicine [29], this study suggested that the director-assessment scores of residents in every EPAs increased significantly over 3-year training, with significant differences between PGY1 vs. PGY2 and PGY1 vs. PGY3 but not PGY2 vs. PGY3. These findings are nearly consistent with previous studies that utilized residency training programs [39] and fellows using the American Board of Pediatrics subspecialty EPAs [16]. The significant differences between

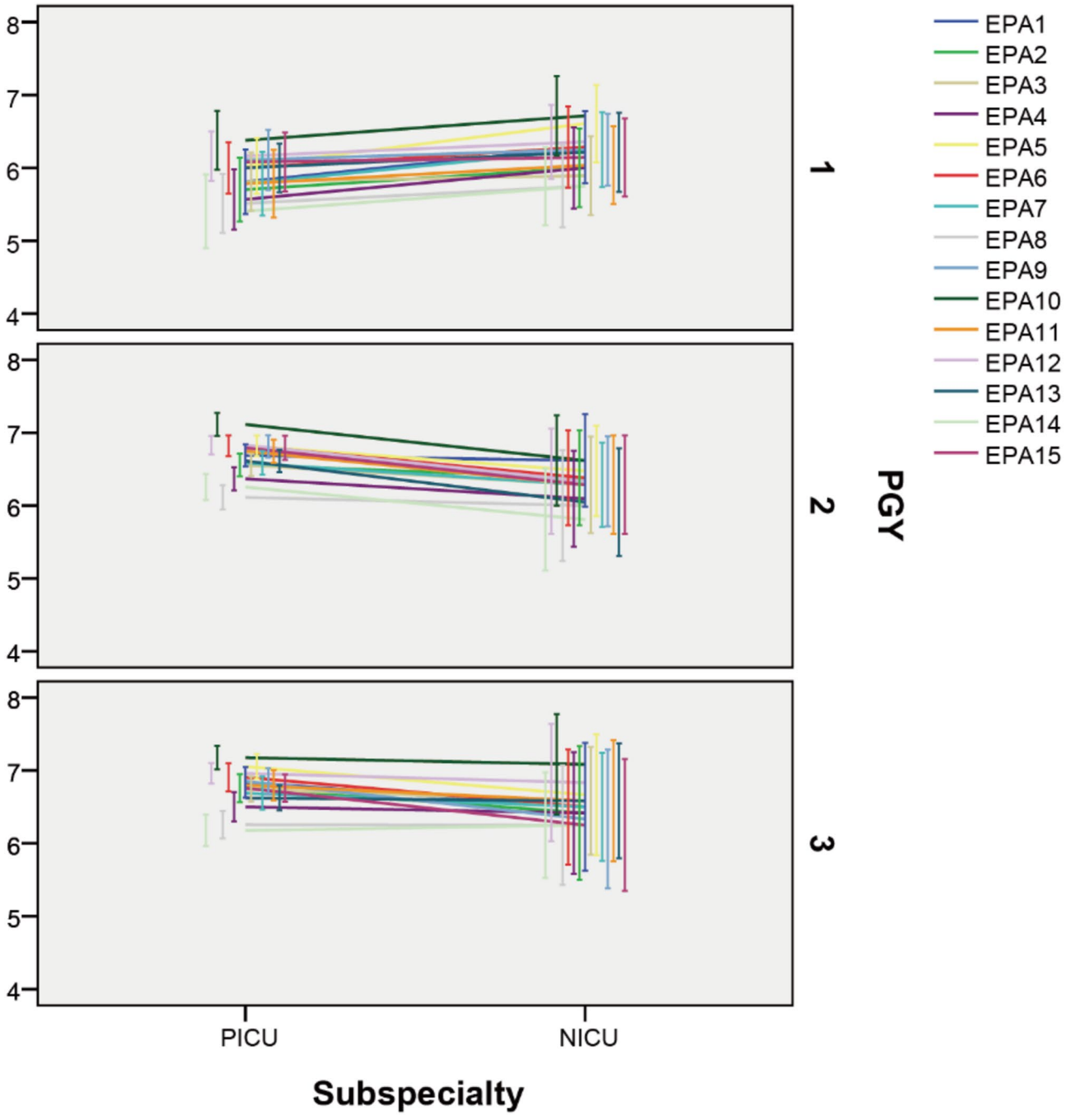

Fig. 4 Line graph of scores of Pediatric Intensive Care Medicine versus Neonatology across different PGYs
PICU: Pediatric Intensive Care Unit(ward of pediatric intensive care medicine); NICU: Neonatal Intensive Care Unit(neonatology ward)

PGY1 and PGY2 demonstrated residents become more experienced and skillful after one year of rotation. However, there was no noticeable improvement after the second year of training, which indicated that there was not enough upgraded training aimed at enhancing residents in PGY2. As upgraded elevation is emphasized in the CBME program, residents are expected to develop their professional medical skills in all EPA categories as their training experience accumulating [40, 41]. It indicates that the curriculum for training residents in these areas requires notable improvement, and directors and regulatory staff should reinforce the idea of upgrading professional teaching in residents between PGY2 and PGY3 years [42, 43]. Residents are expected to meet the standards for unsupervised practice after training in all 15 EPA categories. To achieve it, directors should formulate

Table 7 Scores of self-assessment in different PGYs

EPAs	PGY1	PGY2	PGY3	Chi-square[#]	*P*-value
EPA1	5.5 ± 1.6	5.7 ± 1.5	6.2 ± 1.3	1.951	0.377
EPA2	5.1 ± 1.2	5.5 ± 1.2	6.1 ± 0.9	6.060	***0.048***
EPA3	5.0 ± 1.3	5.2 ± 1.2	5.9 ± 1.0	5.365	0.068
EPA4	4.6 ± 1.3	4.8 ± 1.3	5.7 ± 1.0	6.591	***0.037***
EPA5	6.1 ± 1.3	6.5 ± 1.5	6.7 ± 1.2	3.187	0.203
EPA6	5.8 ± 1.5	6.1 ± 1.4	6.5 ± 1.2	1.973	0.373
EPA7	5.2 ± 1.2	5.6 ± 1.2	6.3 ± 1.0[*]	6.974	***0.031***
EPA8	4.0 ± 1.2	4.8 ± 1.2	5.4 ± 0.9[*]	11.516	***0.003***
EPA9	5.3 ± 1.3	5.4 ± 1.4	6.1 ± 1.2	7.734	0.155
EPA10	6.5 ± 1.1	6.2 ± 1.3	6.6 ± 1.1	1.314	0.518
EPA11	5.0 ± 1.3	5.5 ± 1.3	6.3 ± 1.1[*]	10.622	***0.005***
EPA12	6.2 ± 1.2	6.2 ± 1.4	6.3 ± 1.1	0.050	0.975
EPA13	5.4 ± 1.6	5.3 ± 1.4	5.9 ± 1.0	1.102	0.576
EPA14	4.2 ± 1.7	4.7 ± 1.4	5.4 ± 1.0	5.924	0.052
EPA15	4.0 ± 1.5	4.9 ± 1.5	5.4 ± 1.1[*]	6.949	***0.031***

[#]Kruskal Wallis test. Significant differences are shown in bold($p < 0.05$) and italicized

[*]Mann-Whitney U test revealed a significant difference between PGY1 and PGY3. The Mann-Whitney U test revealed no significant differences between PGY1 and PGY2, PGY2 and PGY3

specific target EPA scores of each PGY. Besides, more specialized trainings such as Mini-CEX.

and more assessments of clinical skills are necessary to implemented in PGY2.

Meanwhile, there were differences across the four residents' positions, except for EPA2(select and interpret auxiliary examinations). Consistent with our previous study on pediatric intensive care medicine, residents of different 4 position varied significantly in most EPA categories. Professional masters, one of the most common positions of these four, had just graduated with a bachelor's degree from a medical school, while permanent staff with doctoral degrees usually had a prolonged research period or more professional knowledge in certain academic fields. In contrast, entrusted training residents and social training residents were more experienced in clinical work and usually had worked for a few years prior to attending standardized residency training, generally with a lesser educational background. Different educational backgrounds and clinical work experience led to different advantages in professional activities, which is urgent for clinical directors to reinforce the personalized training plan for residents in different positions to play to everyone is strengths. The same is the situation in the comparison of EPAs between genders, consistent with previous studies, the minor differences only existed in EPA13(Inform bad news) and EPA15(Manage public health events), in which male residents scored higher than female residents. There were fewer male doctors than female doctors in the pediatric department in most Chinese hospitals, as in our study. The limited number of male residents enrolled in this study, which might lead to inconsistencies between different genders' scores. A further large cohort of residents is required to produce more reliable results. In fact, in daily work, there were no distinct differences between genders in clinical thinking and most medical skills, except for certain medical work requiring great physical strength.

In the GEE model analysis of PGY, gender, and position on EPA scores, all three factors affected the EPA scores to various degrees. PGY had a significant effect on 13 out of 15 EPA scores, while gender affected only four EPA scores significantly, and position affected only three EPA categories. Remarkably, among all the EPAs affected by PGYs, in EPA2(Select and interpret auxiliary examinations), EPA6(Report a case), EPA14(Perform clinical education), and EPA15(Manage public health events), PGY2 ranked the highest with PGY1 ranking the lowest. Residents in PGY2 seemed more skillful and proven than those in PGY1, yet more earnestly and devoted than those in PGY3(the last training year) to some professional activities. As for the four EPAs with a gender gap, the male residents did better than the female ones in EPA8 (recognizing and managing emergent and critical

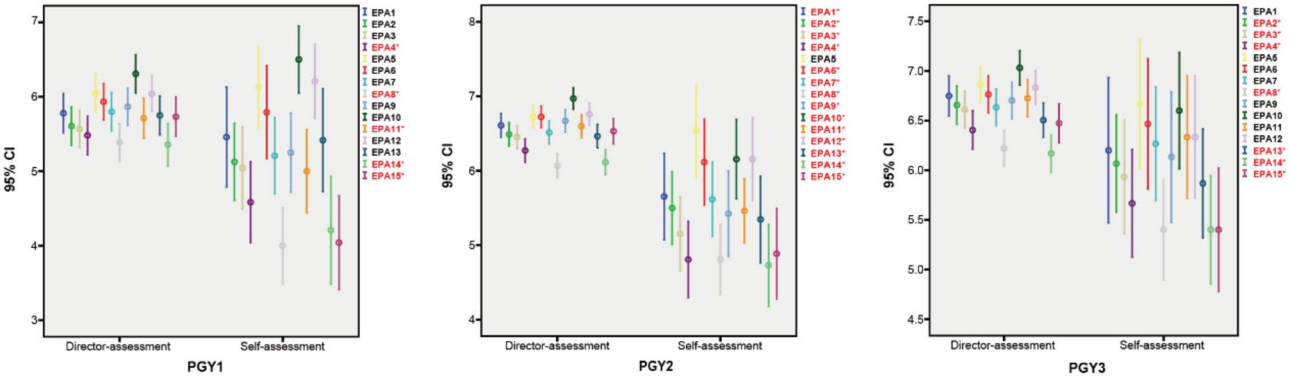

Fig. 5 Error bar chart of director-assessment versus self-assessment within each PGY

[*]Mann-Whitney U test revealed a significant difference in this EPA category between director-assessment scores and self-assessment scores within this PGY. The edges of each bar represent the 95% confidence interval (CI) of the scores in the subgroups

conditions), EPA10 (Perform informed consent), EPA13 (Inform bad news), and EPA14 (Perform clinical education), with EPA14 having the largest score gap(B = 1.040, $p = 0.010$, < 0.05). However, daily medical work related to other EPAs does not show a distinct gender gap. The gender-based difference in outcomes may be due to difference in performance between male and female residents. Male residents may perform better in communicating and presenting with EPA14. Given the current literature and knowledge on bias in evaluation in medical education, we cannot deny that this outcomes may be due to disparity in how faculty assess resident performance [44]. Among the residents with the highest educational background, those in permanent staff positions scored highest in EPA10(Perform informed consent) and EPA13(Inform bad news). Prolonged educational experience helped them preserve more relevant medical knowledge and awareness of the law. Therefore, they could undertake more initiatives to complete medical legal procedures. As the most experienced residents, entrusted training residents had more knowledge from medical practice and were more willing to share with younger physicians, which may be the reason that they did best in EPA14(Perform clinical education).

More than ever, we enrolled pediatric directors from all Pediatric Department subspecialties. Directors from variable subspecialties would judge the residents' behavior in each EPA from the point of view related to their subspecialty features. All EPA categories showed significant differences across different subspecialties. From the perspective of each PGY, the EPA scores of the directors from different subspecialties varied. In PGY1, scores from Pediatric Neurology and Genetic Metabolism Diseases, Pediatric Urology, Endocrinology ranked the highest, while scores from Pediatric Respiratory and Infectious Diseases, Pediatric Gastroenterology, Rheumatology, and Immunology were the lowest ranked. It was presumed that residents in PGY1 had less experience and courage in clinical work, as Neurology and Urology with relatively narrow-spectrum diseases than Respiratory and Gastroenterology, and Novice residents tended to adapt faster in the former subspecialties. It should be noted that there were no scores from Child Healthcare, Pediatric Outpatient, and Emergency Care in PGY1, for the reason that residents of the first training year were considered too young to handle the high-intensity work in Child Healthcare, Pediatric Outpatient, and Emergency Care so that they were set to rotate in this subspecialty until late PGY2 and even until PGY3. When considering PGY2, almost all residents had higher scores, except for Pediatric Gastroenterology, Rheumatology and Immunology, and Pediatric Cardiology. In these two subspecialties, it would require a longer period for residents to improve. However, to varying degrees, the clinical education of directors from these two subspecialties was supposed to be refined and improved. As for PGY3, all scores of EPAs were nearly fully marked, with scores from Child Healthcare, Pediatric Outpatient, and Emergency Care and Pediatric Cardiology ranking the highest and scores from Pediatric Hematology and Oncology, Pediatric Respiratory, and Infectious Diseases ranking the lowest. It cannot be ignored that there was no obvious improvement in the scores of each EPA from PGY1 to PGY3 in Pediatric Hematology and Oncology, Pediatric Respiratory and Infectious Diseases. This warned us to focus more on step-up teaching in clinical education in the aforementioned subspecialties. As there are limited studies on pediatric subspecialty implementation [45], more samples and multi-centered studies are needed for further verification in this field. In brief, educators and regulatory agencies would need to implement EPA-based assessments more broadly or efficiently in pediatric subspecialties, as suggested previously [46].

For the comparison of scores from Pediatric Intensive Care Medicine and Neonatology, EPA9(transfer and hand over a patient), EPA11(perform basic operation), and EPA15(manage public health events) showed significant differences between the two subspecialties, with PICU ranking higher. There would be differences in clinical skill training procedures between these two subspecialties because of the differences in the patient groups. Residents needed more time and skill to be qualified to perform certain work independently on patients in the Neonatology Department than in the PICU. As for EPA11(perform basic operation), it is much more technically difficult to perform basic surgery in newborns considering their low weight and immature organ development. Furthermore, the puncture location the depth of lumbar puncture and bone marrow puncture are different in newborns and children, which may also affect the puncture performance of residents.

Interestingly, with respect to self-assessment scores, only some EPA scores were significantly different across PGYs and between each of the two PGYs. There was no obvious improvement in the self-assessment scores between PGY1 and PGY2 and PGY2 and PGY3. Even in PGY1 versus PGY3, only some EPAs(EPA7 (recognize and manage general clinical conditions), EPA8 (recognize and manage emergent and critical conditions), EPA11 (perform basic operation), and EPA15 (manage public health events)) showed significant differences. This situation was inconsistent with scores from the director-assessment, especially in PGY2 and PGY3, with director-assessment scoring higher than self-assessment. Regarding the differences between director- and self-assessment, we hold the view that the director-assessment is more objective and reliable. Because the directors were professors in each subspecialty and they exhibited

a higher degree of mastery for entrustable levels of each EPA. It can also be proved by the evidence that the vast majority of previous EPAs studies focused on the director-assessment. This result indicated that the residents were not very clear about their real strength and were not confident enough to judge themselves. Moreover, there is an urgent need for an efficient feedback program between directors and residents so that they can note their advantages and disadvantages in time for continuous improvement.

Our study has several strengths. It reported the implementation and feasibility of EPAs in the Chinese standardized training of residents in pediatric and different pediatric subspecialty departments. It established significant differences in EPA performance within residents from lower PGY to higher PGY and provided a well-structured framework to guide residents in the development of clinical knowledge, skills, and attitudes. The results may have some feedback effect on the residency training. According to the unsatisfying improvement between PGY2 and PGY3, more clinical trainings in each EPA categories were urgently needed in the second and the third years. We analyzed the effects of PGY, gender, and resident position on EPAs scores, confirming that PGY, gender, and position were correlated with EPA scores to various degrees. In this study, we also analyzed EPA scores from different subspecialties, indicating that some subspecialties require more efficient step-up training procedures. Incongruity between director-assessed and self-assessed scores indicates the need for an efficient feedback program for residents' continuous improvement.

There are also several limitations to our study. First, the findings reflect the experience at only a single center. Second, this was a cross-sectional study that enrolled residents trained in pediatrics within the last three years. Hence this study could not show changes in EPA scores from PGY1 to PGY3 for a single resident. A longitudinal study may be valid and a multicenter longitudinal study would be of great value.

In summary, our study shows that EPA assessments had a certain discriminating efficacy in Chinese standardized residency training across training years in the pediatric department, with scores rising with the PGY year. PGY, gender, and resident position affected EPA scores. Some differences in director-assessment from different subspecialties warned directors from related subspecialties to focus on step-up clinical training. Given the incongruities between the resident-assessed and director-assessed scores, an efficient feedback program is needed for residents' continuous improvement.

Conclusion

The study findings suggest that EPA assessments is feasible among different PGYs in standardized Chinese residency training in pediatric medicine and its subspecialties. Postgraduate year had a significant impact on EPA scores, while gender and resident position also affected EPA scores to a certain extent. Improved stratified teaching programs are required for better subspecialty consistency.

Abbreviations

EPAs	Entrustable Professional Activities
PGYs	postgraduate years
PGME	postgraduate medical education
CBME	competency-based medical education
Mini-CEX	mini-clinical-evaluation exercises
DOPS	direct observation of procedural skills
SOAP	subjective-objective-assessment-plan
GEE	generalized estimated equation

Acknowledgements

We are grateful to all staff at the Medical Training Office at Qilu Hospital, Shandong University. Finally, we are especially grateful to the residents and directors whose data were included in this study.

Author contributions

Dr. Yun Zhang initiated the study, participated in the design and coordination, performed statistical analysis, and drafted the manuscript. Xiaoyu Jiao collected the original data and performed part of the statistical analyses. Dongxiu Zhao and Professor Yuankai Zhang guided the surveys and manage the standard procedures.Tong Yue and Ke Yang conducted all the questionaires among residents and directors, along with all the follow-up work. Professor Cuifen Zhao and Professor Aijun Zhang initiated the study and edited the manuscript. All authors approved the final manuscript as submitted and agreed to be accountable for all aspects of the work.

Funding

This research was a key Qilu Medical Characteristics Undergraduate Education and Teaching Research Project of Shandong University(qlyxjy-202301) and an Undergraduate Education and Teaching Reform Research Project of Shandong University (2023Y038). All phases of this study were supported by a grant from the Innovative Research Project on Standardized Residency Training in the Qilu Hospital of Shandong University (ZPZX2023B03) and Shandong Provincial Natural Sciences Foundation(ZR2023QC160).

Data availability

We declare that the datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

We solemnly state the consent that was obtained from all of the participants was informed. This study adhered to the Declaration of Helsinki and passed the review of the Qilu Hospital of Shandong University's ethics committee(KYLL-202205-011).

Consent for publication

This study has obtained the consent for publication of all the participants.

Competing interests

The authors declare no competing interests.

Data sharing statement

Individual participant data (including data dictionaries) will be made available in addition to study protocols, the statistical analysis plan, and the informed consent form. The data will be made available upon publication to researchers who provide a methodologically sound proposal for achieving the goals of

the approved proposal. These proposals should be submitted to yunzhang@email.sdu.edu.cn.

Received: 15 October 2024 / Accepted: 25 April 2025
Published online: 12 May 2025

References

1. Ten CO, Scheele F. Competency-based postgraduate training: can we Bridge the gap between theory and clinical practice? ACAD MED. 2007;82:542–7.
2. Carraccio C, Englander R, Gilhooly J, et al. Building a framework of entrustable professional activities, supported by competencies and milestones, to Bridge the educational continuum. ACAD MED. 2017;92:324–30.
3. Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system–rationale and benefits. N Engl J Med. 2012;366:1051–6.
4. Ten CO. Entrustability of professional activities and competency-based training. MED EDUC. 2005;39:1176–7.
5. Ten CO. Nuts and bolts of entrustable professional activities. J Grad Med Educ. 2013;5:157–8.
6. Moore D, Young CJ, Hong J. Implementing entrustable professional activities: the yellow brick road towards competency-based training? ANZ J SURG. 2017;87:1001–5.
7. Boyce P, Spratt C, Davies M, McEvoy P. Using entrustable professional activities to guide curriculum development in psychiatry training. BMC MED EDUC. 2011;11:96.
8. Mulder H, Ten CO, Daalder R, Berkvens J. Building a competency-based workplace curriculum around entrustable professional activities: the case of physician assistant training. MED TEACH. 2010;32:e453–9.
9. Warm EJ, Held JD, Hellmann M, et al. Entrusting observable practice activities and milestones over the 36 months of an internal medicine residency. ACAD MED. 2016;91:1398–405.
10. O'Dowd E, Lydon S, O'Connor P, Madden C, Byrne D. A systematic review of 7 years of research on entrustable professional activities in graduate medical education, 2011–2018. MED EDUC. 2019;53:234–49.
11. Peters H, Holzhausen Y, Boscardin C, Ten CO, Chen HC. Twelve tips for the implementation of EPAs for assessment and entrustment decisions. MED TEACH. 2017;39:802–7.
12. Deitte LA, Gordon LL, Zimmerman RD, et al. Entrustable professional activities: ten things radiologists do. ACAD RADIOL. 2016;23:374–81.
13. McCloskey CB, Domen RE, Conran RM, et al. Entrustable professional activities for pathology: recommendations from the college of American pathologists graduate medical education committee. Acad Pathol. 2017;4:1531714411.
14. Shaughnessy AF, Sparks J, Cohen-Osher M, Goodell KH, Sawin GL, Gravel JJ. Entrustable professional activities in family medicine. J Grad Med Educ. 2013;5:112–8.
15. Young JQ, Hasser C, Hung EK, et al. Developing End-of-Training entrustable professional activities for psychiatry: results and methodological lessons. ACAD MED. 2018;93:1048–54.
16. Mink RB, Schwartz A, Herman BE, et al. Validity of level of supervision scales for assessing pediatric fellows on the common pediatric subspecialty entrustable professional activities. ACAD MED. 2018;93:283–91.
17. Rose S, Fix OK, Shah BJ, Jones TN, Szyjkowski RD. Entrustable professional activities for gastroenterology fellowship training. Gastroenterology. 2014;147:233–42.
18. Powell DE, Wallschlaeger A. Making sense of the milestones: entrustable professional activities for pathology. HUM PATHOL. 2017;62:8–12.
19. Parker TA, Guiton G, Jones MJ. Choosing entrustable professional activities for neonatology: a Delphi study. J PERINATOL. 2017;37:1335–40.
20. Myers J, Krueger P, Webster F, et al. Development and validation of a set of palliative medicine entrustable professional activities: findings from a mixed methods study. J PALLIAT MED. 2015;18:682–90.
21. Kerth JL, Bosse HM. The future of postgraduate training?? An update on the use of entrustable professional activities in pediatric professional training?. ACAD PEDIATR. 2024;24:1035–7.
22. Mink RB, Carraccio CL, Herman BE, et al. Subspecialty pediatrics investigator network. Relationship between Epa level of supervision with their associated subcompetency milestone levels in pediatric fellow assessment. BMC MED EDUC. 2023;23:720.
23. Mink RB, Schwartz A, Herman BE, et al. Steering committee of the subspecialty pediatrics investigator network (SPIN). Validity of level of supervision scales for assessing pediatric fellows on the common pediatric subspecialty entrustable professional activities. ACAD MED. 2018;93:283–91.
24. Weiss PG, Schwartz A, Carraccio CL, et al. Achieving entrustable professional activities during fellowship. Pediatrics. 2021;148:e2021050196.
25. Langhan ML, Boyer DL, Hsu D, et al. Implementing entrustable professional activities in pediatric fellowships: facilitating the process. Pediatrics. 2024;153:e2023065024.
26. Turner DA, Schwartz A, Carraccio C, et al. Subspecialty pediatrics investigator network (SPIN). Continued supervision for the common pediatric subspecialty entrustable professional activities May be needed following fellowship graduation. ACAD MED. 2021;96:S22–8.
27. Pitts S, Schwartz A, Carraccio CL, et al. Fellow entrustment for the common pediatric subspecialty entrustable professional activities across subspecialties. ACAD PEDIATR. 2022;22:881–6.
28. Li S, Qi X, Li H, et al. Exploration of validity evidence for core residency entrustable professional activities in Chinese pediatric residency. FRONT MED. 2024;10:1301356.
29. Yun Z, Jing L, Junfei C, et al. Entrustable professional activities for Chinese standardized residency training in pediatric intensive care medicine. FRONT PEDIATR. 2022;10:919481.
30. Schumacher DJ, West DC, Schwartz A, Association of Pediatric Program Directors Longitudinal Educational Assessment Research Network General Pediatrics Entrustable Professional Activities Study Group, et al. Longitudinal assessment of resident performance using entrustable professional activities. JAMA NETW OPEN. 2020;3:e1919316.
31. Ten CO, Chen HC, Hoff RG, Peters H, Bok H, van der Schaaf M. Curriculum development for the workplace using entrustable professional activities (EPAs): AMEE guide 99. MED TEACH. 2015;37:983–1002.
32. Ten CO, Hart D, Ankel F, et al. Entrustment decision making in clinical training. ACAD MED. 2016;91:191–8.
33. Kerth JL, van Treel L, Bosse HM. The use of entrustable professional activities in pediatric postgraduate medical education: A systematic review. ACAD PEDIATR. 2022;22:21–8.
34. Bramley AL, McKenna L. Entrustable professional activities in entry-level health professional education: A scoping review. MED EDUC. 2021;55:1011–32.
35. Meyer EG, Chen HC, Uijtdehaage S, Durning SJ, Maggio LA. Scoping review of entrustable professional activities in undergraduate medical education. ACAD MED. 2019;94:1040–9.
36. O Connor P, Lydon S, O Dowd E, Byrne D. Entrustable professional activities for early health professionals: what's next? MED EDUC. 2021;55:984–6.
37. Amiel JM, Andriole DA, Biskobing DM, et al. Revisiting the core entrustable professional activities for entering residency. ACAD MED. 2021;96:S14–21.
38. Weiss PG, Schwartz A, Carraccio CL et al. Achieving entrustable professional activities during fellowship. Pediatrics. 2021;148.
39. Schumacher DJ, West DC, Schwartz A, et al. Longitudinal assessment of resident performance using entrustable professional activities. JAMA Netw Open. 2020;3:e1919316.
40. Schultz K, Griffiths J. Implementing Competency-Based medical education in a postgraduate family medicine residency training program: A Stepwise approach, facilitating factors, and processes or steps that would have been helpful. ACAD MED. 2016;91:685–9.
41. Li ST, Tancredi DJ, Schwartz A, et al. Competent for unsupervised practice: use of pediatric residency training milestones to assess readiness. ACAD MED. 2017;92:385–93.
42. McMillan JA, Land MJ, Leslie LK. Pediatric residency education and the behavioral and mental health crisis: A call to action. Pediatrics. 2017;139.
43. Peter NG, Forke CM, Ginsburg KR, Schwarz DF. Transition from pediatric to adult care: internists' perspectives. Pediatrics. 2009;123:417–23.
44. Klein R, Julian KA, Snyder ED, et al. From the gender equity in medicine (GEM) workgroup. gender Bias in resident assessment in graduate medical education: review of the literature. J GEN INTERN MED. 2019;34:712–9.
45. Kerth J, van Treel L, Bosse HM. The use of entrustable professional activities in pediatric postgraduate medical education: A systematic review. ACAD PEDIATR. 2022;22:21–8.

46. Taylor DR, Park YS, Smith CA, Karpinski J, Coke W, Tekian A. Creating entrustable professional activities to assess internal medicine residents in training: A Mixed-Methods approach. ANN INTERN MED. 2018;168:724–9.

Publisher's note