

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect

Biomedical Journal

journal homepage: [www.elsevier.com/locate/bj](http://www.elsevier.com/locate/bj)

## Original Article

# Comparing learning outcomes among postgraduate year trainee groups



Peng-Wei Hsu <sup>a,e,\*</sup>, Ren-Huei Fu <sup>b,d,e</sup>, Yu-Che Chang <sup>c,d,e</sup>

<sup>a</sup> Department of Neurosurgery, Chang Gung Memorial Hospital at Linkou, Taoyuan, Taiwan, ROC

<sup>b</sup> Department of Pediatrics, Chang Gung Memorial Hospital at Linkou, Taoyuan, Taiwan, ROC

<sup>c</sup> Department of Emergency Medicine, Chang Gung Memorial Hospital at Linkou, Taoyuan, Taiwan, ROC

<sup>d</sup> Department of Medical Education, Chang Gung Memorial Hospital at Linkou, Taoyuan, Taiwan, ROC

<sup>e</sup> College of Medicine, Chang Gung University, Taoyuan, Taiwan

## ARTICLE INFO

## Article history:

Received 18 October 2018

Accepted 4 October 2019

Available online 22 October 2020

## Keywords:

Competence

Multiple-choice questions

Objective structured clinical examination

Postgraduate training

## ABSTRACT

**Background:** The objective of postgraduate year (PGY) training programs is to inculcate in medical graduates the expected levels of skills in patient care. This study compared the core clinical competencies of trainees who received PGY training at Chang Gung Memorial Hospital by attending the pilot training program in different groups.

**Methods:** We used six 10-min test stations for clinical performance evaluation, which comprised four and two test stations designed for objective structured clinical examination and procedural skill, respectively, to evaluate the learning outcomes of the trainees. The trainees were divided into three groups according to the training programs that they had attended.

**Results:** The aspects of clinical performance included history taking, physical examination, medical communication, logical thinking, and problem-solving abilities. The trainees who selected the surgery-based training program exhibited a higher performance at the station for aseptic surgical preparation than the other two groups ( $p = 0.0261$ ). The trainees who selected the internal medical training program ( $p = 0.0020$ ) exhibited a higher performance at the station for abdominal pain in children.

**Conclusions:** A well-designed postgraduate training program should develop trainees' competencies, particularly clinical operational skills. The results of this study may provide useful insight into methods for improving the design of training programs. Additional investigation is necessary for understanding the effects of different programs on the clinical performance of trainees.

\* Corresponding author. Department of Neurosurgery, Chang Gung Memorial Hospital at Linkou, 5, Fusing St, Gueishan, Taoyuan 333, Taiwan.

E-mail address: [leon3096@gmail.com](mailto:leon3096@gmail.com) (P.-W. Hsu).

Peer review under responsibility of Chang Gung University.

<https://doi.org/10.1016/j.bj.2019.10.002>

2319-4170/© 2019 Chang Gung University. Publishing services by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### At a glance commentary

#### Scientific background on the subject

We use structured objective structured clinical examinations (OSCEs) to evaluate the core competencies and professional performance of PGY trainees in clinical practice. The reliability and validity analysis of the OSCEs and statistical analysis were used to verify the results of the study.

#### What this study adds to the field

Taiwan's medical education system change to 2-years post-graduate training gradually since year 2011. The pilot program was launched in year 2015 and there were no studies analyzing the outcomes. The results of present study provide valuable information that can be used to improve the design of medical training programs.

Previously, medical students in Taiwan acquired clinical skills and cultivated patient care skills through knowledge-centered learning. They began clinical work in professional subdivisions immediately after graduating, consequently, an overall lack of abilities was perceived in general clinical medical care [1]. The United Kingdom designed a postgraduate general education training system in 1951 [2], and the United States developed the postgraduate year (PGY) comprehensive medical training course in 1970 [3]. Both training programs were intended to solve the problem of the lack of comprehensive clinical care abilities among medical graduates. Since then, advanced countries have established comprehensive medical training programs for 1 or 2 years to improve the patient care abilities, clinical skills, and physician–patient communication skills of medical graduates.

After the severe acute respiratory syndrome (SARS) pandemic in 2003, the postgraduate training program for general medicine was implemented by the Taiwanese government to address the need for improvement in the professional training provided to medical graduates. After detailed and rigorous planning, the Taiwan Joint Commission on Hospital Accreditation developed a PGY training program to improve the abilities of medical graduates. The PGY trainees were expected to develop their skills and abilities in patient-centered care after completing general medical training, including medical knowledge, clinical skills, and professional attitudes. The PGY training program, which was a 3-month course, was launched in Taiwan in 2003. In 2011, a 1-year training program was launched and will eventually be transformed into a 2-year course in 2019. Finally, Taiwan's medical education system will change from seven years of medical education plus one year of post-graduate training to 6-year medical education plus 2-years post-graduate training. It will be a big change in the medical education system.

In the first year of the proposed 2-year training course, participants are expected to receive comprehensive medical training to enhance their overall skills in general medical care.

Currently, a 1-year training program is being implemented. In the second year, specialized training will be provided depending on the individual interests and requirements of the trainees. In August 2013, Chang Gung Memorial Hospital (CGMH) launched a pilot project for the second year PGY training program.

Objective structured clinical examination (OSCE), a useful tool, is usually used to assess the learning result of medical trainees. A well designed OSCE can be used to assess the core competencies in medical graduates. Since there were no studies comparing the outcomes of different pilot training programs and core competence of the trainees, we thought that an effective assessment is necessary. Therefore, the aim of our study is to evaluate the core competencies by well-designed assessments. By using OSCE, this study compared the differences in the core clinical competencies of trainees who received PGY training at CGMH by attending the second year training courses in different groups. The results of this study will provide further understand about the impact of different education program on learning outcomes. We also expect that the quantitative and qualitative information in the learning resulted between the different training groups facilitate improvements in the design of the program and accurately assess the implementation of comprehensive medical training in Taiwan.

### Methods

Concerning about the research ethics, the study is approved by the Institutional Review Board (IRB No. 103-7377B).

Between August 2015 and July 2017, 134 PGY trainees completed the training program at CGMH one year after they graduated from the medical school. The trainees were assessed using six 10-min test stations designed for clinical performance evaluation in the last month of one-year training program. Four stations were designed for OSCE and were related to internal medicine, surgery, obstetrics and gynecology, and pediatrics, whereas the remaining two stations were designed for assessing clinical skills.

Depending on the training program attended, the trainees were divided into the following three groups:

Program 1 was a conventional training program (currently provided in Taiwan) designed for training in general medicine. The training program comprised 3 months of internal medicine training, 2 months of surgery training, 1 month of training in pediatrics, 1 month of obstetrics and gynecology, 1 month of an elective course, and 4 months of emergency department and community medical training.

Program 2 and Program 3 was designed as a pilot project for the second year course of the 2-year training program.

Program 2 (surgical training program) was designed for training in surgery and the obstetrics and gynecology subspecialties. The program comprised 6 months of training on surgery and gynecology, 1 month of internal medicine training, 2 months of an elective course, and 3 months of emergency medicine and community medicine.

Program 3 (internal medicine training program) was designed for training in the internal medicine and the pediatrics subspecialties. The program comprised 6 months of

training on internal medicine and pediatrics, 1 month of surgery training, 2 months of an elective course, and 3 months of training on emergency medicine and community medicine.

Twelve trainees, selected randomly from each group ( $n = 36$ ), participated in the clinical performance evaluation. There were 8 male and 4 female in the group 1 (conventional training program), 7 male and 5 female in the group 2 (surgical training program), and 7 male and 5 female in the group 3 (internal medicine training program). All the participating trainees were the post-graduate year-one residents. The criteria for passing or failing at each test station were determined using the Angoff method. A 3-point Likert scale was used to evaluate the result of each item on the checklist, which was defined as not completed (0), partially completed (1), and completed (2). The final score obtained at each test station was determined using the following formula: (score obtained/maximum obtainable score)  $\times$  100. The mean score was then calculated across all test stations. All the raters were qualified by the Taiwan Association of Medical Education and had completed the rater-training program.

An unpaired *t* test, analysis of variance, and the Wilcoxon rank sum test were used to analyze the data in SPSS Version 19.0 (SPSS Inc., Chicago, IL, USA).  $p < 0.05$  indicated statistical significance.

## Results

The examination topics designed for the assessment of learning outcomes in internal medicine, surgery, gynecology, and pediatrics were palpitations, burn injuries, uterine fibroids, and abdominal pain in children. Several aspects of performance were evaluated at the different test stations. History taking, physical examination, and medical communication were evaluated at the surgery and internal medicine test stations. At the stations for gynecology and pediatrics, the clinical reasoning, problem-solving abilities, physical examination skills, and medical communication of the trainees were assessed. The reliability (Cronbach's alpha coefficient, as known as Cronbach's  $\alpha$ ) and validity (Kendall's coefficient of concordance, as known as Kendall's  $\omega$ ) of the test stations were analyzed before the assessment [Table 1]. An example of the checklist of the test station (palpitation) was attached on Table 2. The test stations that were designed for the evaluation of clinical skills assessed the trainees' skills at aseptic surgical preparation and use of infection-protective clothing.

The analysis and comparison results of trainee performance at individual test stations revealed that the trainees who selected the surgical training program exhibited a significantly higher performance ( $p = 0.0261$ , Wilcoxon rank sum test) at the test station for aseptic surgical preparation than the other two groups. The trainees who selected the internal medical training program exhibited a higher performance ( $p = 0.0020$ , Wilcoxon rank sum test) at the station for abdominal pain in children. The trainees' skills with burn injuries, uterine fibroids, palpitations, and the use of infection-protective clothing did not differ significantly among the three groups [ $p = 0.8090, 0.3829, 0.3855, \text{ and } 0.6597$ , respectively, Table 3].

## Discussion

The development of Taiwan's comprehensive medical graduate training program can be divided into four stages. The training duration extended from 3 months to 6 months and eventually one year. The training program contents extended from general and community medicine to more comprehensive internal medicine and surgery training. Other specialties such as emergency medicine, pediatrics, and obstetrics/gynecology were also added. The objective of this stage was to enable every medical graduate to develop sufficient comprehensive abilities in clinical care. Immediately after 2019, this training program will enter the fourth stage, which is a 2-year training program. The first year of the training course will follow the third-stage program. After entering the second year of the program, trainees will be allowed to select a training system according to their interests. The trainees can then decide whether to become surgeons, physicians or other specialists. In August 2013, CGMH launched a pilot trial for the second year of the 2-year PGY training program.

According to the rules of the Accreditation Council for Graduate Medical Education (ACGME), the PGY training program is expected to highlight and develop six core competencies in medical graduates [4]. The program uses various tools for the effective and credible assessment of trainee learning outcomes and for improving training program content by using the results of the assessment. These assessments include quantitative assessments (e.g. Mini-Clinical Evaluation Exercise (Mini-CEX), Direct Observation of Procedural Skills (DOPS), Case-based Discussion (CbD)) and qualitative, reflective writing sections (e.g. Medical Ethics and Legislation Report, Medical Care Quality Report and Personal

**Table 1** Blueprint of assessment aspects, reliability and validity at different test stations for objective structured clinical examination.

Station	Topic	Contents	Reliability (Cronbach's $\alpha$ )	Validity (Kendall's $\omega$ )
Surgery	Burn injury	History taking, physical examination, medical communication	0.968	0.871
Internal Medicine	Palpitation	History taking, physical examination, medical communication	0.884	0.857
Gynecology	Uterine fibroids	Clinical reasoning and problem solving, physical examination, medical communication	0.851	0.807
Pediatric	Abdominal pain	Clinical reasoning and problem solving, physical examination, medical communication	0.911	0.825

**Table 2 Example of checklist; Station: Palpitation.**

Score	0	1	2
Item	Not Completed	Partially Completed	Completed
<b>Communication skills (medical communication):</b>			
1. Identify the patient and self-introduction with good manner			
2. Look at the patient and listen during conversation			
3. Understand the chief complaint of the patient			
4. Explain with easy language, not medical terminology			
<b>History taking:</b>			
5. Duration of episode			
6. Frequency of episode			
7. Explore inducing factors and relieving factors			
8. Combined symptoms with episode			
9. Find current medication			
10. Allergy history			
11. Past medical history			
12. Family history			
<b>Physical examination and procedures:</b>			
13. Arrange EKG exam			
14. Reading of EKG (Atrial fibrillation with RVR 及 Normal sinus rhythm)			
15. Check vital signs (heart rate, blood pressure)			
<b>Communication skills (medical communication):</b>			
16. Explain the diagnosis and future risks			
17. Suggestion of future plan or treatment			

The gray shade indicates that this option cannot be selected in this checklist.

Development Report). According to Taiwanese regulations for e-portfolios, PGY trainees are expected to fill the e-portfolios numerous times over the course of their training course.

However, these assessments are all workplace-based assessments evaluated by individual clinical teachers. There were no standardized assessments to evaluate the PGY trainees'

**Table 3 Differences in performance among the three groups of trainees.**

Station	Program, Score/mean (SD)			<sup>a</sup> p value
	1, General medicine	2, Surgery, and Obstetrics/Gynecology	3, Internal medicine and Pediatrics	
Palpitation	81.37 (7.45)	77.94 (13.88)	85.29 (7.73)	0.3855
Burn injury	59.80 (11.59)	57.11 (16.52)	57.84 (11.38)	0.8090
Uterine fibroids	64.58 (11.77)	59.58 (8.58)	63.96 (10.03)	0.3829
Abdominal pain in children	67.82 (9.29)	66.67 (7.49)	75.23 (12.50)	0.0020
Infection-protective clothing	66.90 (10.82)	69.68 (9.21)	66.44 (9.29)	0.6597
Aseptic surgical preparation	62.91 (26.44)	68.20 (16.32)	50.22 (11.36)	0.0261

<sup>a</sup> Wilcoxon rank sum test.

training program and outcomes. OSCEs with standardized patients (SPs) and checklists can be used to assess interpersonal communication skills as well as patient care. Multiple-choice questions (MCQ) and oral tests are useful tools for assessing medical knowledge, whereas OSCEs and checklists effectively assess professionalism. OSCEs, SPs, checklists, and MCQ tests can be used to assess trainees' abilities to learn and improve in practice. They are also useful for evaluating system-based practices [5]. Considering their proven efficacy, MCQ tests, OSCE with SPs, and checklist were used to analyze the learning outcomes of the different training programs included in this study.

MCQ tests have been used for departmental or comprehensive exams to determine whether progress has been achieved or as a certification for learning outcomes [6–9]. OSCE with interactive situations involving SPs can be tailored to meet specific educational goals, and students' performance can be reliably evaluated [10,11]. According to the literature, if the number of sites and trainees is sufficiently high, then the reliability of the assessment is as high as 0.90 [12]. The specific skills that we assessed using OSCE were medical history taking, physical examination, interpersonal communication, technical skills, data interpretation, differential diagnosis, and treatment decisions. Trainees' technical skills can be objectively and structurally assessed using an evaluation list that has been validated for its reliability and effectiveness [13,14]. In this study, the Cronbach's  $\alpha$  for reliability is ranged from 0.851 to 0.968, the Kendall's  $\omega$  for validity is from 0.807 to 0.871 [Table 1].

In the current study, to compare the performance of different groups of trainees under the same test conditions, we used the same assessment blueprint while designing the topics of assessment. For example, in the test stations for heart palpitations, burn injuries, uterine fibroids, and abdominal pain in children, the tests included medical history inquiry, physical examination, and medical communication. To assess the effectiveness of gynecology and pediatric learning, the physical examination skills, logical thinking, problem-solving abilities, and medical communication skills of the trainees were assessed. After data analysis, we found that the trainees who had attended Program 3 exhibited a significantly higher performance at the test station for abdominal pain in children than the other groups ( $p = 0.0020$ ). However, the same results were not observed at the test station for palpitations ( $p = 0.2207$ ) and infection-protective clothing ( $p = 0.6597$ ). The trainees who had attended Program 2 exhibited a significantly higher performance in aseptic surgical preparation than the other trainees ( $p = 0.0261$ ). However, they did not exhibit a higher performance than the other trainees did at the test stations for burn injuries ( $p = 0.8090$ ) and uterine fibroids ( $p = 0.3829$ ).

The reason for these results may be related to the training plan arrangement and the design of the station. There is a lack of general medical training in the present pilot training program. Training in Pediatrics was limited in Program 1 and Program 2 (1 month in Program 1 and not included in Program 2). The clinical presentations of diseases in children differ from those in adults. These two factors caused the relatively low performance in Pediatrics

of the trainees who selected Program 1 and 2. Program 2 provided the trainees with more opportunities to enter the operating room to learn the related techniques than did Program 1 and Program 3; consequently, the trainees in Program 2 exhibited a higher performance in preparation in a sterile operating room than the trainees in other programs. By contrast, trainees who selected Programs 1 and 3 generally had fewer opportunities to wear infection-protective clothing than those who selected Program 2; therefore, the performance of the two groups is lower than did in Program 2, although there was no significant statistical difference. Further researches with OSCE stations containing similar assessment aspects and recruitment of more PGY participants are necessary to certify our assumptions.

---

## Conclusions

The results of this study show that highly specialized training may result in insufficient training for trainees in other fields (for example, in the pediatrics training in this study). Therefore, in the 2-year training program, general medical training is necessary in the first year. It enables medical graduates to develop their general medical care abilities. Specialized training in the second year can strengthen the trainees' technical skills and prepare them for specialized field training in the future. Combining the first year of general medical training and the second year of specialist training, the trainees could access a comprehensive ability in clinical care. A limitation of the present study is the relatively low number of test stations and trainees involved in the clinical performance assessment process. However, the results provide valuable information that can be used to improve the design of training programs.

---

## Conflicts of Interest

Peng-Wei Hsu received partial funding for this study from Chang Gung Memorial Hospital (grant No. CDRPG3E0031). The authors have no conflicts of interest to declare.

---

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bj.2019.10.002>.

---

## REFERENCES

- [1] Kuhlmann TP, Fang WL, Fan Y. Physicians' views on how specialty-specific the first year of residency should be. *Acad Med* 1991;66:237–9.
- [2] Cohen M. The GP preregistration house officer: the potential learning experience of primary care. *Hosp Med* 1998;59:502–4.
- [3] Cassie JM, Armbruster JS, Bowmer MI, Leach DC. Accreditation of postgraduate medical education in the

- United States and Canada: a comparison of two systems. *Med Educ* 1999;33:493–8.
- [4] Leach DC. The ACGME competencies: substance or form? Accreditation Council for Graduate Medical Education. *J Am Coll Surg* 2001;192:396–8.
- [5] Yang YY, Lee FY, Hsu HC, Huang CC, Chen JW, Lee WS, et al. A core competence-based objective structured clinical examination (OSCE) in evaluation of clinical performance of postgraduate year-1 (PGY1) residents. *J Chin Med Assoc* 2011;74:198–204.
- [6] Anastakis DJ, Wanzel KR, Brown MH, McIlroy JH, Hamstra SJ, Ali J, et al. Evaluating the effectiveness of a 2-year curriculum in a surgical skills center. *Am J Surg* 2003;185:378–85.
- [7] Anastakis DJ, Cohen R, Reznick RK. The structured oral examination as a method for assessing surgical residents. *Am J Surg* 1991;162:67–70.
- [8] De Champlain AF, Melnick D, Scoles P, Subhiyah R, Holtzman K, Swanson D, et al. Assessing medical students' clinical sciences knowledge in France: a collaboration between the NBME and a consortium of French medical schools. *Acad Med* 2003;78:509–17.
- [9] Hsu PW, Hsieh MJ, Fu RH, Huang JL, Liao MC, Lee ST. Comparing the outcomes of different postgraduate year training programs in Taiwan. *Biomed J* 2015;38:544–9.
- [10] Hall-Turner WJ. An experimental assessment carried out in an undergraduate general practice teaching course (OSCE examination). *Med Educ* 1983;17:112–9.
- [11] Ortega P, Park YS, Girotti JA. Evaluation of a medical Spanish elective for senior medical students: improving outcomes through OSCE assessments. *Med Sci Educ* 2017;27:329–37.
- [12] Hodges B, Regehr G, Hanson M, McNaughton N. Validation of an objective structured clinical examination in psychiatry. *Acad Med* 1998;73:910–2.
- [13] De La Garza JR, Schmidt MW, Kowalewski KF, Benner L, Müller PC, Kenngott HG, et al. Does rating with a checklist improve the effect of E-learning for cognitive and practical skills in bariatric surgery? A rater-blinded, randomized-controlled trial. *Surg Endosc* 2019;33:1532–43.
- [14] Lockyer J, Singhal N, Fidler H, Weiner G, Aziz K, Curran V. The development and testing of a performance checklist to assess neonatal resuscitation megacode skill. *Pediatrics* 2006;118:e1739–44.