

Design of an instrument to evaluate entrustable professional activities for general surgery

Diseño de un instrumento para evaluar las actividades profesionales confiables en cirugía general

Carlos R. Cervantes-Sánchez^{1,2*}, Haydeé Parra-Acosta², and Juan C. Cantú-Reyes²

¹Servicio de Cirugía General, Hospital General de Chihuahua Dr. Salvador Zubirán Anchondo; ²Departamento de Investigación Educativa, Facultad de Medicina y Ciencias Biomédicas, Universidad Autónoma de Chihuahua. Chihuahua, Chihuahua, México

Abstract

Objective: Competency-based medical education has become increasingly prevalent among surgical accreditation bodies that conduct assessments to ensure the competency of future professionals and judge the adequacy of training programs. **Materials and Methods:** Cross-sectional study to validate an instrument to evaluate entrustable professional activities (EPA) according to the Single Program of Medical Specialties through content and construct validity by expert judgment, and internal consistency with Cronbach's alpha. **Results:** Four of the rubrics scored high to optimal for univocity and relevance. Only one rubric had to be modified in two of its items. Internal consistency was high (Cronbach's alpha of 0.97). When applying the instrument, the residents showed differences in the level of competence according to postgraduate year, identifying better performance in the last year. **Discussion:** The instrument was validated with a high degree of reliability in univocity and relevance, and high internal consistency. The residents demonstrated appropriate EPA to manage surgical patients safely, efficiently, with quality and warmth. The ability to reach the strategic level of performance is the main component of the competency-based medical education through the EPA.

Keywords: Competency-based education. Competency-based evaluation. Surgical education. Validity study.

Resumen

Objetivo: La educación médica basada en competencias se ha vuelto cada vez más frecuente entre los organismos de acreditación quirúrgica que realizan evaluaciones para asegurar la competencia de los futuros profesionales y juzgar la idoneidad de los programas de formación. **Material y Métodos:** Estudio transversal para validar un instrumento para evaluar las actividades profesionales confiables (APC) según el Programa Único de Especialidades Médicas mediante validez de contenido y constructo por juicio de expertos, y consistencia interna con alfa de Cronbach. **Resultados:** Cuatro de las rúbricas obtuvieron nivel alto a óptimo para univocidad y relevancia. Solo una rúbrica debió ser modificada en dos de sus ítems. La consistencia interna fue alta (alfa de Cronbach de 0.97). Al aplicar el instrumento, los residentes mostraron diferencias en el nivel competencial según el año de posgrado, identificando mejor desempeño el último año. **Discusión:** El instrumento fue validado con un alto grado de fiabilidad en univocidad y relevancia, y una alta consistencia interna. Los residentes demostraron poseer las APC necesarias para tratar pacientes quirúrgicos de manera segura, eficiente, con calidad y calidez. La capacidad para alcanzar el nivel estratégico de desempeño es el componente principal de la educación médica basada en competencias a través de las APC.

Palabras clave: Educación basada en competencias. Evaluación basada en competencias. Educación quirúrgica. Estudio de validez.

*Correspondence:

Carlos R. Cervantes-Sánchez
E-mail: ccervantes@uach.mx

Date of reception: 17-01-2022

Date of acceptance: 06-04-2022

DOI: 10.24875/CIRUE.M22000493

Cir Cir (Eng). 2022;90(6):798-806

Contents available at PubMed

www.cirugiaycirujanos.com

2444-0507/© 2022 Academia Mexicana de Cirugía. Published by Permanyer. This is an open access article under the terms of the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Higher education is under continuous scrutiny as performance standards increase progressively. According to American educational associations, such as the American Association for Higher Education, the American Education Research Association, the American Psychological Association and the National Council on Measurement in Education, learning assessment is a scientific process aimed at determining what can be known regarding performance skills, and how they are better measured in terms of validity, accuracy, reliability, analysis and reporting¹.

The frameworks for skill-based medical education have become more and more frequent among surgical education accreditation entities, including the Consejo Mexicano de Cirugía General. Traditional medical residency training programs use a model based on time, expecting students to achieve the skills at the end of the program by continuous exposure during the years of training. In contrast, students in skill-based medical education programs must show their skills in a variety of activities and contexts throughout their programs in order to successfully complete their training. The feasibility of performing periodic skill assessments is still a challenge, especially for surgical skills, due to time constraints in the operating room, safety concerns, and the reluctance of supervising personnel to allow students to perform surgery independently².

However, there is no single definition of skill that encompasses all the important domains of the professional medical practice and that is universally accepted. In response to this, the Accreditation Council for Graduate Medical Education determined six areas that defined professional skills as the habitual and judicious use of communication, knowledge, technical abilities, clinical reasoning, emotions, values, and day-to-day reflexion for the benefit of the individual and community being served³.

It is known that advanced performance in the operating room depends on the combination of several sets of abilities, for which any objective evaluation must be adapted and thoroughly tested⁴.

In Mexico, the medical specialties curriculum has been designed and implemented since 1994 through the Plan Único de Especialidades Médicas (PUEM; Single Medical Specialties Plan), defined as a high-quality medical practice based on three key areas of professional expertise: medical care, research development, and educational activities⁵.

The Consejo Mexicano de Cirugía General (Mexican Committee for General Surgery) has acknowledged that, for many years, certification and re-certification has been based on written and oral exams, and this has put in question whether accredited surgeons have the necessary skills for clinical care of patients, and whether the training systems of the country are adequately acceptable⁶. In the last few decades, both the Royal College of Physicians and Surgeons of Canada and the American Board of Medical Specialties have implemented certification programs to ensure that physicians are committed to a permanent learning and skill assessment process. However, relevance, administrative requirements, and evidence of effectiveness have been increasingly questioned⁷.

Since 2001, the Accreditation Council for Graduate Medical Education has unsuccessfully implemented several projects (*Outcomes*, *Next*), with the purpose of increasing the quality of graduate medical education. The concept of entrustable professional activities was introduced in the Netherlands in 2005, as a practical construct to describe the work of a physician according to their skills. For this reason, as part of a larger approach to medical education based on skills, the surgical specialty has led the way to introduce them to graduate surgical specialization⁸.

For this reason, in 2016, the American Board of Medical Specialties considered the entrustable professional activities as a potential model to improve the evaluation and feedback of surgical apprentices. However, its effectiveness in evaluation the achievement of expertise and autonomy is unknown⁹.

Currently, the assessment of surgical skill is still the weakest aspect of the evaluation, and therefore, it is imperative to have an objective assessment strategy for knowledge, skills, and attitudes of the specialized physician through a Integral Evaluation Model of Clinical and Surgical Skills, not only in a transversal manner, but with a design that allows to assess the pros and cons of the method, by means of a research-action process, with the possibility of correction and improvement, adapting to the need of both participants and the society according to the time and circumstances. Therefore, it is essential to have a tool to perform an integral assessment of the entrustable professional activities of the specialty of general surgery.

Method

A validity and reliability study was performed in order to design a tool comprised of five forms with levels of

Table 1. Definition of univocity levels

Definition of the univocity scale levels		
Univocity	Meaning	Assigned value
Optimal univocity (OU)	The item may be unequivocally understood and interpreted in one single way.	4
High Univocity (HU)	The item may be subject to interpretation but can be mainly or generally understood in only one way.	3
Low Univocity (LU)	The item may be understood in different ways and is closer to ambiguity.	2
No Univocity (NU)	The item may not be understood or may be interpreted with different meanings, falling into ambiguity.	1

Table 2. Definition of relevance levels

Definition of the relevance scale levels		
Relevance levels	Meaning	Assigned value
Optimal relevance (OR)	The item may belong to unequivocally belong to the construct to which it is directed	4
High Relevance (HR)	The item may belong mainly to the construct to which it is directed	3
Low Relevance (LR)	The item may belong to several constructs or may not be very relevant to the construct to which it is directed	2
No Relevance (NR)	The item may not belong to the construct or it is not relevant to the construct to which it is directed	1

performance in order to evaluate the entrustable professional activities of the residents of a general surgery program in the class of 2020-2021 in a second-degree care public hospital, through the following phases.

Tool design

The department of Educational Research of the Faculty of Medicine and Biomedical Sciences of Universidad Autónoma de Chihuahua developed five forms according to the PUEM, in terms of: 1-3) clinical-surgical attention (preoperative, transoperative and postoperative management); 4) search and transfer of knowledge, and 5) research skills. Each area was comprised of 6 or 7 items, classified in four levels of performance: a) receptive; b) decisive; c) autonomous, and d) strategic¹⁰.

Expert judgment

For content validity the expert judgment technique was employed¹¹ based on two criteria: univocity and relevance, on a estimation scale of four levels, from “optimal univocity” to “no univocity” (Table 1), with the participation of 12 experts (six general surgeons, three clinical researchers, and three medical education researchers) who received the forms via email, according to their field of expertise. For the relevance criteria, a scale with meanings and levels similar to those of univocity was employed (Table 2).

For the forms that assess clinical-surgical attention, general surgery specialists with a minimum of 5 years of experience were considered experts. For the category that assesses research skills, researchers belonging to the National Researchers System were petitioned. Lastly, for the category of educational skills,

Table 3. Interpretation of univocity and relevance rates

Univocity rate (Ur)		Relevance rate (Rr)	
nOU	Number of response obtained from the optimal univocity level	nOR	Number of response obtained from the optimal relevance level
nHU	Number of response obtained from the high univocity level	nHR	Number of response obtained from the high relevance level
nLU	Number of response obtained from the low univocity level	nLR	Number of response obtained from the low relevance level
nNU	Number of response obtained from the no univocity level	nNR	Number of response obtained from the no relevance level
nTU	Number of total responses obtained from the univocity level	nTR	Number of total responses obtained from the relevance level
VOU	Value assigned to the optimal univocity level	VOR	Value assigned to the optimal relevance level
VHU	Value assigned to the high univocity level	VHR	Value assigned to the high relevance level
VLU	Value assigned to the low univocity level	VLR	Value assigned to the low relevance level
VNU	Value assigned to the no univocity level	NNR	Value assigned to the no relevance level
VMU	Maximum value of the univocity scale levels	VMR	Maximum value of the relevance scale levels

$$i_U = \frac{(\sum n_{UO} \cdot V_{UO}) + (\sum n_{UE} \cdot V_{UE}) + (\sum n_{UB} \cdot V_{UB}) + (\sum n_{UN} \cdot V_{UN})}{\sum n_{TU} \cdot V_{MU}}$$

$$i_P = \frac{(\sum n_{PO} \cdot V_{PO}) + (\sum n_{PE} \cdot V_{PE}) + (\sum n_{PB} \cdot V_{PB}) + (\sum n_{PN} \cdot V_{PN})}{\sum n_{TP} \cdot V_{MP}}$$

Figure 1. Formula to obtain univocity and relevance indexes. Adapted from: Carrera FX et al.¹¹

physicians with graduate degrees in medical education were considered.

The defined procedure to obtain univocity and relevance rates (Table 3) allows us to ponder expert responses as described by Carrera¹², where, using the formula proposed in Figure 1¹¹ it was possible to obtain both rates with a maximum value of 4 and a minimum of 1 for each item of the validation tool, setting certain criteria to determine the preservation (≥ 0.80), modification (≤ 0.79) or elimination (≤ 0.59) of each assessed item (Table 4).

Internal consistency

Once the tools have been validated, they were applied to resident doctors of the surgery service of Hospital General de Chihuahua, "Dr. Salvador Zubirán Anchondo" in order to be observed and assessed while performing

their everyday activities (outpatient visits, hospital rooms, classroom) (Fig. 2) by the instructors of the course (heteroevaluation) with the use of electronic forms (Google forms). In order to assess the internal consistency of the tools, Cronbach alpha was employed¹³.

The evaluation of the preoperative management was performed through six categories of skills: 1) structured clinical interview; 2) physical exploration; 3) diagnostic process; 4) surgical management proposal; 5) informing the patient, and 6) doctor-patient relationship.

The evaluation of the transoperative management was approached through the following seven areas: 1) placement of the patient; 2) preparation of the surgical area; 3) washing and dressing for the surgical act; 4) performance within the surgical team; 5) performance of the surgical technique; 6) closure of surgical wound, and 7) conduct during the surgical act.

The evaluation of the postoperative management was through six areas of expertise: 1) clinical interview (focused on recovery); 2) physical exploration (evaluating recovery); 3) identification and resolution of complications; 4) proposal of postoperative management until recovery; 5) informing the patient of their progress, and 6) doctor-patient relationship after surgery.

The search for and transfer of knowledge was approached through six areas of expertise: 1) critical thinking; 2) creativity; 3) writing skills; 4) speaking skills; 5) reading comprehension, and 6) computer skills.

Table 4. Univocity and relevance criteria

Criteria according iU		Criteria according to iR	
$iU \geq 0.80$	Items are kept in their original form	$iR \geq 0.80$	Items are kept in their original form.
$iU \leq 0.79 \geq 0.60$	Items are modified in drafting.	$iR \leq 0.79 \geq 0.60$	Items are modified in drafting or location.
$iU \leq 0.59$	Items are removed.	$iR \leq 0.59$	Items are removed.

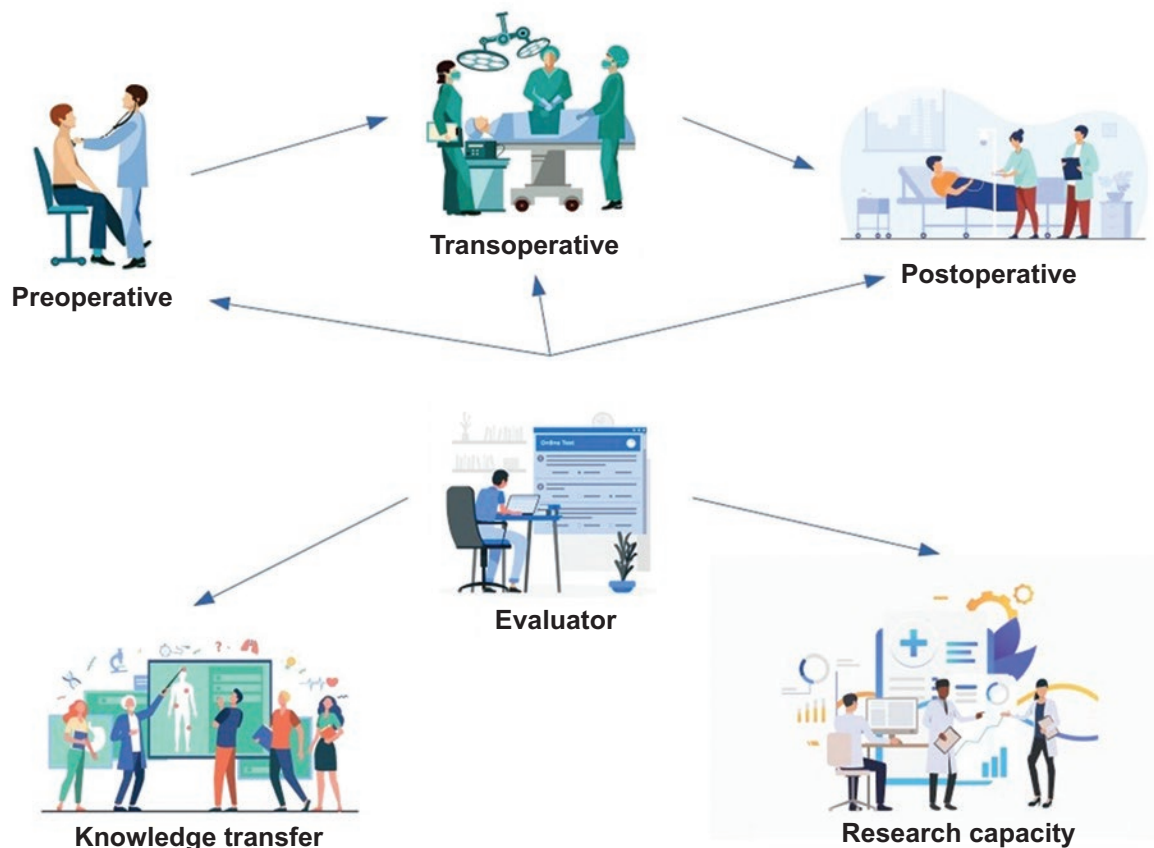


Figure 2. Diagram of reliable professional activities evaluation. (Own creation with images from macrovector_official, pch.vector, katemangostar and storles, obtained from <https://www.freepik.com>).

Research skills were approached through seven areas of expertise assessed: 1) generating a research inquiry; 2) systematic search for information; 3) scientific methodology aimed at answering the research question; 4) observes, analyzes, and describes the medical phenomenon; 5) analysis, discussion, and drafting of the results, and 7) creation and presentation of a research report.

The procedures described meet the ethical standards of the Committee of Research Ethics of Subdirección de Enseñanza e Investigación del Hospital General de Chihuahua "Dr. Salvador Zubirán Anchondo"

(CEI-0189-2017) and the Committee of Research Ethics of Secretaría de Investigación y Posgrado de la Facultad de Medicina y Ciencias Biomédicas de la Universidad Autónoma de Chihuahua (CI-006-21), in accordance with the 1975 Declaration of Helsinki and subsequent revisions (including the 64th Asamblea General de Fortaleza, Brazil, in October of 2013)¹⁴. Both committees are registered with the National Bioethics Committee of Secretaría de Salud. The project is registered in ClinicalTrials.gov (NCT05179980). In all cases, informed consent was obtained from residents and health care personnel involved (nurses), and patients.

Table 5. Skill: research skill (initial)

Key	Mastery of skill to assess	Univocity level	Relevance level
CI1	It identifies the needs of the global context consistent with the challenges of the knowledge society	0.9	0.95
CI2	Analyzes and retrieves relevant information	0.9	0.95
CI3	Applies methodological processes	0.85	0.95
CI4	Shows observation ability	0.85	0.85
CI5	Adopts an ethical attitude when processing information	0.75	0.95
CI6	Generates new knowledge that contributes to problem solving	0.75	1.0
CI7	Develops different types of scientific communication	0.9	1.0

Table 6. Skill: research skill

Key	Mastery of skill to assess	Univocity level	Relevance level
CI1	It identifies the needs of the global context consistent with the challenges of the knowledge society	0.9	0.95
CI2	Analyzes and retrieves relevant information	0.9	0.95
CI3	Applies methodological processes	0.85	0.95
CI4	Shows observation ability	0.85	0.85
CI5	Adopts an ethical attitude when processing information	0.85	0.95
CI6	Generates new knowledge that contributes to problem solving	0.85	1.0
CI7	Develops different types of scientific communication	0.9	1.0

Results

Regarding, the validity of the contents of the forms by means of expert judgment, it was seen that, in the form for the research skills, items 5) *adopts an ethical attitude when processing information*, and 6) *generates new knowledge to contribute to problem solving*, showed

values $\leq 0.79 \geq 0.60$, and therefore, they were able to be modified (Table 5). Once modified, it was returned to the same evaluators after a washout period of 1 month, obtaining values of > 0.85 , and therefore, they are considered entrustable after the correction (Table 6).

In terms of the forms that evaluate skills related to clinical and surgical care (preoperative, trans-operative, and postoperative), and the one that allows to evaluate the search for and transfer of knowledge, it was observed that the areas assessed retained their original wording as they showed univocity and relevance of > 0.80 (Table 7-10).

The assessment of the internal consistency for each form obtained a Cronbach's alpha value of 0.97, with a Guttman split-half coefficients of 0.86, which is considered highly acceptable as it is very close to 1.0, which indicates that the tool is highly entrustable¹³.

A pilot test was conducted to test the tool with the residents of the program offered by the Medicine and Biomedical Sciences Faculty of Universidad Autónoma de Chihuahua, distributed by degree in the following manner: six first-year (R1), four second-year (R2), three third-year (R3), and two fourth-year (R4) residents. The age range was between 25 and 31 years (mean: 28.26; standard deviation: 1.9), and 4 (26.6%) were females and 11 (73.4%) males; the distribution of females was two in R1, one in R2, and one in R4.

The assessments were performed by four physicians who are members of the service, the main professor of the program and the assistant, the research project leader, and the head of the service.

In the entrustable professional activities related to clinical and surgical care (preoperative, transoperative, and postoperative management), clear differences were found in the performance between the different degrees, whereby the R1 group tends to be between the receptive and resolute levels, R2 between the resolute and autonomous levels, R3 stayed at the level of autonomy, and finally R4 were at the autonomous level and reached the strategic level.

In the entrustable professional activities of search for and transfer of knowledge an almost homogeneous trend is seen between the R1 and R2 groups at the receptive and resolute, while the autonomous level is reached starting on group R3, and the strategic level is only reached by the R4 group.

However, in the entrustable professional activities that assess research skills, all residents showed a receptive level, although one member of the R1 group obtained a higher score, and only a female member of the R4 group reached the autonomous level.

Table 7. Skill: preoperative management

Key	Mastery of skill to assess	Univocity level	Relevance level
MPrO1	Applies structured interview	0.94	0.88
MPrO2	Examines patient completely and systematically	0.97	0.97
MPrO3	Identifies the presence or lack of surgical pathology	0.97	1.0
MPrO4	Proposes and explains the specific surgical treatment	0.94	0.88
MPrO5	Informs patients and family members of condition	0.94	0.94
MPrO6	Nurtures and promotes the doctor-patient relationship	0.94	0.94

Table 8. Skill: transoperative management

Key	Mastery of skill to assess	Univocity level	Relevance level
MT1	Ensures the position of the patient and the lighting are adequate	1.0	1.0
MT2	Ensures that the preparation of the surgical area is adequate, and knows the theoretical rationale	0.9	0.93
MT3	Cleans and dresses themselves at the right time and knows the right technique	1.0	0.96
MT4	I capable of performing all the positions of the surgical team, and to initiate the approach	1.0	0.93
MT5	Is able to perform the different repair techniques of the surgical condition	0.96	0.9
MT6	Closes the surgical wound by layers	0.96	0.96
MT7	Their attitude in the entire surgical act is one of mastery, respect, and leadership towards the well-being of the patient	0.9	0.87

Discussion

The main objective of this study was to design a tool to assess, in the most objective way possible, the entrustable professional activities associated with the specialty of general surgery of the Faculty of Medicine and Biomedical Sciences of Universidad Autónoma de

Table 9. Skill: postoperative management

Key	Mastery of skill to assess	Univocity level	Relevance level
MPOp1	Applies structured interview	0.94	0.94
MPOp2	Examines patient completely and systematically	1.0	1.0
MPOp3	Identifies and resolves the presence or lack of complications	0.97	0.91
MPOp4	Proposes and explains specific medical and/or surgical treatment	0.97	0.97
MPOp5	Informs patients and family members of recovery	1.0	1.0
MPOp6	Nurtures and promotes the doctor-patient relationship	0.91	0.91

Table 10. Skill: search for and transfer of knowledge

Key	Mastery of skill to assess	Univocity level	Relevance level
BTC1	Critical thinking	0.95	0.91
BTC2	Creativity	0.83	0.91
BTC3	Writing	0.91	0.91
BTC4	Speaking	0.83	0.83
BTC5	Reading comprehension	0.87	0.87
BTC6	Computer skills	0.83	0.91

Chihuahua, defined based on PUEM. With this in mind, five forms with descriptors for complex variables was designed with the areas of the entrustable professional activities through the interactions of residents with patients, their families, and the healthcare team in the context of the actual clinical care of the patient, that is, the top of Miller's pyramid. Thus, this is a contributing factor for the evaluation to be considered an ongoing learning and reflexion strategy for the physician in training¹⁵.

Beyond measuring the performance or lack thereof of a certain skill or ability, the tools designed allow for the identification of performance level, which is invaluable for students to be aware of their own situation at any given time, and allows them to identify areas of opportunity for improvement, which turns this method into a training process that fosters learning and helps improve the academic and professional performance of the specialist in training.

Another advantage of the tool is that, due to its design, it can be used in a 360° evaluation environment, as it can be used by the instructors of the course, healthcare personnel of the institution (nurses, social workers, etc.) and other residents of the institution and other specialties, and it can also be used as a self-assessment and co-assessment tool, in addition to heteroevaluation.

This evaluation allowed the identification of differences between the level of expertise according to academic degree, showing better performance in higher stages, and in this way, they showed the necessary entrustable professional activities to treat patients with conditions that require surgical intervention in a safe and effective manner, with quality and warmth, which is consistent with the findings of other researchers^{16,17}.

It is worth noting that the entrustable professional activities that generally showed a higher level of expertise were defined by PUEM as relative to the assistance activity, proving to possess sufficient capacity for the clinical care of the patients from the preoperative to the postoperative period. This allows us to identify some skills where the residents seem insecure or not experienced, obtained lower scores, while in some skills they appear as secure and competent as residents in higher grades.

A similar skill pattern to the assistance activity can be seen in the entrustable professional activities related to the search for and transfer of knowledge, in the sense that, it is evident that, in the last year, residents have more ease and effectiveness both in creating and presenting classes and review topics. However, in the entrustable professional activities related to research skills, there is a cause for concern, as the majority of the residents show a complete lack of competence in topics related to research, which leaves an opportunity niche for improvement, starting on the first year of residency.

The evaluation of this cycle will allow to compare it against subsequent evaluations (annuals or biannual even), making this practice an integral learning assessment, as each evaluation is accompanied by individual feedback for each resident, in order to allow the apprentice to know their level of performance with the purpose of identifying strengths and weaknesses to correct their habitual day-to-day clinical activity.

In conclusion, the design of this tool has been validated with a high degree of reliability, both in terms of univocity and relevance, and high internal consistency. This validation process provides access to a complete and objective evaluation method to evaluate the entrustable professional activities of surgery residents during

their graduate program training, which is the next step for the residency evaluation, far from being summative, to become a constructivist learning method, providing residents with valuable information to identify skills where they need to reach a higher level of sophistication through a wider professional practice.

Acknowledgements

The authors acknowledge the valuable collaboration of Dr. Mario Humberto Cardiel Ríos, Dr. Rubén Burgos Vargas, Dr. Abel Jalife Montaño, Dr. María de Lourdes de la Torre Díaz, Dr. Gloria Eugenia Queipo García, Dr. Jesús Arturo Medina Ramírez, Dr. Pedro Alexandro Márquez, Dr. Laura Elizabeth Piñón Gaytán, Dr. Ubaldo García, and Dr. César Aragón Quintana, who selflessly agreed to act as expert judges during the validation process for the content of the tools assessed.

Funding

The authors declare that they have received no funding.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Ethical disclosures

Protection of people and animals. The authors state that the procedures followed the ethical guidelines of the responsible human experimentation committee and were performed according to the World Medical Association and the Declaration of Helsinki.

Data confidentiality. The authors declare that they have followed their center's protocols on the publication of patient data.

Right to privacy and informed consent. The authors have obtained the informed consent of the patients and/or subjects referred to in the article. This document is in the possession of the corresponding author.

References

1. Parker PE, Fleming PD, Beyerlein S, Apple D, Krumsieg K. Differentiating assessment from evaluation as continuous improvement tools [for engineering education]. 31st Annual Frontiers in Education Conference. Impact on Engineering and Science Education. Conference Proceedings (Cat. No.01CH37193). Available from: <https://www.semanticscholar.org/paper/Differentiating-assessment-from-evaluation-as-tools-Parker-Fleming/37418c016324131c379337030bab06b29430fea4>

2. Fahim C, Wagner N, Nousiainen MT, Sonnada R. Assessment of technical skills competence in the operating room: a systematic and scoping review. *Acad Med*. 2018;93:794-808.
3. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA*. 2002;287:226-35.
4. Azari D, Greenberg C, Pugh C, Wiegmann D, Radwin R. In search of characterizing surgical skill. *J Surg Ed*. 2019;76:1348-63.
5. Plan Único de Especializaciones Médicas (PUEM) en Cirugía General. Facultad de Medicina, UNAM, División de Estudios de Posgrado, Subdivisión de Especializaciones Médicas, Coordinación de Desarrollo Curricular. México, D.F.; 2009. Plan de estudios aprobado por el Consejo Universitario el 21 de abril de 1994. Coordinación editorial del Dr. Bernardo Pintos Aguilar. Ciudad de México; junio de 2016.
6. Anaya-Prado R, Ortega-León LH, Ramírez-Solis ME, Vázquez-García JA, Medina-Portillo JB, Ayala-López EA. Evaluación de competencias quirúrgicas. Estudio piloto mexicano. *Cir Cir*. 2012;80:261-9.
7. Price D, Campbell C, Van Hoof TJ, ElChamaa R, Jeong D, Chappell K, et al. Definitions of physician certification used in the North American literature: a scoping review. *JCEHP*. 2020;40:147-57.
8. Brasel KJ, Klingensmith ME, Englander R, Grambau M, Buyske J, Sarosi G, et al. Entrustable professional activities in general surgery: development and implementation. *J Surg Educ*. 2019;76:1174-86.
9. Albright JB, Meier AH, Ruangvoravat L, VanderMeer TJ. Association between entrustable professional activities and milestones evaluations: real-time assessments correlate with semiannual reviews. *J Surg Educ*. 2020;77:e220-8.
10. Tobón S. Formación integral y competencias. Pensamiento complejo, currículo, didáctica y evaluación. 4.ª ed. Bogotá, Colombia: Ecoe;2013.
11. Carrera FX, Vaquero E, Balsells MA. Instrumento de evaluación de competencias digitales para adolescentes en riesgo social. *EDUTEC, Revista Electrónica de Tecnología Educativa*. 2011;(35). Available from: <http://edutec.rediris.es/revalec2/revalec35>.
12. Carrera Farran FX. Uso de los diagramas de flujo y sus efectos en la enseñanza aprendizaje de contenidos procedimentales. Área de tecnología (E.S.O.). *Revista Interuniversitaria de Formación del Profesorado*. 2005;19:223-5.
13. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. *Res Sci Educ*. 2018;48:1273-96.
14. World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. *JAMA*. 2013;310:2191-4.
15. Champin D. Evaluación por competencias en la educación médica. *Revista Peruana de Medicina Experimental y Salud Pública*. 2014;31:566-71.
16. Kouzmina E, Mann S, Chaplin T, Zevin B. An evaluation of the surgical foundations curriculum: a national study. *J Surg Educ*. 2021;78:914-26.
17. Dwyer T, Wadey V, Archibald D, Kraemer W, Slade Shantz J, Townley J, et al. Cognitive and psychomotor entrustable professional activities: can simulators help assess competency in trainees? *Clin Orthop Relat Res*. 2016;474:926-34.