

Access this article online
Quick Response Code:

Website: www.jehp.net
DOI: 10.4103/jehp.jehp_44_23

Validation of a form for assessing the professional performance of residents in cardiology by nurses

German Busch, María V. Rodríguez Borda¹, Pablo I. Morales¹, Mónica Weiss¹, Graciana Ciambrone², Juan P. Costabel, Eduardo Durante³, Ricardo Gelpi⁴, Alberto E. Alves De Lima

Abstract:

BACKGROUND: There is a strong need for transformation in our assessment systems from one that evaluates performance based on levels of training to another that focuses on professional competence to meet the expected requirements for the practice of the profession. The aim of this study is to validate for the first time a Spanish version of a new tool for assessing the professional performance of residents by nurses newly developed in the Ottawa Hospital (O-RON).

MATERIALS AND METHODS: After the author's written authorization, the original O-RON form was translated and cross-culturally adapted. Then we conducted a prospective observational study in two cardiology centers in the city of Buenos Aires. The validity of the tools was evaluated by the ability of the instrument to discriminate the level of experience of the residents according to their post-graduate year level. Data is expressed as percentages and frequencies of the qualifications obtained in the different questions. The chi-square test was used to assess the significance of the differences obtained. A generalizability test was used to evaluate reliability. Feasibility was defined as a minimum of 4 assessments per resident per evaluation round. Satisfaction of evaluators was assessed using a survey with a 10-point scale designed by the authors.

RESULTS: A total of 838 evaluations were performed. Regarding validity, the 15-item form could significantly discriminate the experience of the residents according to their postgraduate year level ($P < 0,005$). Thirty evaluations per resident are required to obtain reliable results. The tool is feasible to implement and an average of 4.55 assessment per resident per evaluation round were achieved throughout the entire experience. This value remained stable during the 8 rounds (1st: 4.65; 2nd: 4.34; 3rd: 4.47; 4th: 6.17; 5th: 4.56; 6th: 4.08; 7th: 4.36; 8th: 3.91). The levels of satisfaction among the evaluators were acceptable.

CONCLUSION: The Spanish version of the O-RON form can provide residents with a valuable source of feedback from the eyes of nurses on important aspects of their professional training. This tool, positively assessed by the raters, significantly discriminates residents' experience. Its implementation is feasible in our environment, and it is user-friendly, though it requires a considerable number of assessments to achieve high reliability.

Keywords:

Cardiology, feedback, nurses, observation

Introduction

There is a strong need for transformation in our assessment systems from one

that evaluates performance based on levels of training to another that focuses on professional competence to meet the expected requirements for the practice of

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

How to cite this article: Busch G, Rodríguez Borda MV, Morales PI, Weiss M, Ciambrone G, Costabel JP, *et al.* Validation of a form for assessing the professional performance of residents in cardiology by nurses. *J Edu Health Promot* 2023;12:127.

Department of
Clinical Cardiology,
ICBA—Instituto
Cardiovascular, Buenos
Aires, Argentina,
¹Nursing Department,
ICBA—Instituto
Cardiovascular, Buenos
Aires, Argentina,
²Department of Cardiology,
Hospital Alemán De
Buenos Aires, Argentina,
³Department of Medicine,
Hospital Italiano De
Buenos Aires, Argentina,
⁴National Scientific and
Technological Research
Council (CONICET),
Buenos Aires, Argentina

Address for correspondence:

Dr. German Busch,
Department of
Clinical Cardiology,
ICBA—Instituto
Cardiovascular, Av.
Del Libertador
6302, (C1428DCO),
Buenos Aires, Argentina.
E-mail: gbusch@icba.com.ar

Received: 11-01-2023
Accepted: 14-02-2023
Published: 28-04-2023

the profession. Workplace-based assessments (WPBAs) are instruments used in the actual setting of the practice and are considered one of the best methods of assessing professional competence.^[1] There is a strong need to develop evaluation tools that are capable of reflecting workplace performance.^[2,3]

Most WPBAs are based on direct observation of residents (Rs) by physician supervisors (PS). A PS observes and evaluates an R while he/she obtains medical history and performs a physical examination in an inpatient or outpatient setting. After asking the R for a diagnosis and treatment plan, the PS completes a short evaluation form and gives direct feedback about the R's performance.^[3,4] This tool has some limitations: on the one hand, it requires at least 7 to 11 individual evaluations to draw generalizable conclusions about the competencies of each R,^[4-6] and on the other hand, there is evidence that performance under the observation of their supervisors may not reflect their performance in independent, unobserved situations. Several studies have demonstrated that when physicians are "observed," they alter their diagnostic and prescribing behaviors and demonstrate better compliance with hand hygiene protocols and clinical practice recommendations.^[7]

Multiple interactions with other healthcare professionals and patients occur in daily practice, where key processes like communication are not observed by PS, as when a nurse asks the R about a patient's problem or medication.^[8] A comprehensive assessment of the R's performance requires tools capable of evaluating him/her from a different perspective than that of the PS. Assessing an R through observation while he/she interacts with other healthcare professionals without the direct observation of physicians provides a more authentic assessment of his/her performance in real-world settings.

In most clinical contexts, nurses are the healthcare professionals that typically interact most with Rs. Nurses directly interact with Rs on a routine basis, providing an excellent opportunity to observe them on several competencies such as communication or professionalism. The nature of a nurse's work demands close interaction with the patient and other healthcare workers and can provide a different and more important perspective on an R's skills, especially in terms of interpersonal communication and teamwork. Having nurses' assessments of an R's performance may provide a more complete picture of his/her performance.

Recently, Dudek *et al.*^[9] published the psychometric results of the Ottawa Resident Observation Form for Nurses (O-RON). This instrument is a WPBA tool used

by nurses to evaluate the performance of different Rs with a 15-item form rated on a 3-point Likert scale, one global judgment yes/no question regarding whether they would want the R on their team, and a space for written comments. The results showed that the O-RON demonstrates promise as a work-place based assessment tool to provide residents and training programs with feedback on aspects of their performance in a hospital ward through the eyes of the nurses. The instrument was easy to use and had solid evidence for validity and reproducibility.

The aim of this study is to document if a Spanish version of the O-RON form implemented in a wide range of clinical settings and in different cardiology residency programs will achieve adequate levels of validity, reliability, feasibility, and rates of satisfaction among nurses like those achieved in its original language.

Material and Methods

Study design and setting

We conducted a prospective observational study in two cardiology centers in the city of Buenos Aires (ICBA—*Instituto Cardiovascular* and *Hospital Alemán*) with the intention of validating a Spanish version of the O-RON form.

Study participants and sampling

Participants included cardiology residents of the two participating centers who had completed at least one year of their formation program. Nurses were invited to participate anonymously and voluntarily in the validation of a new assessment tool for residents' performance if they have worked in common areas for at least a 30-day period before initiating this protocol.

Data collection and technique

After the author gave her written authorization, the original O-RON form was translated and cross-culturally adapted [Figure 1].^[10]

The tool was administered in eight observation rounds for 9 months (September 2021–May 2022). Each observation period lasted between 4 and 6 weeks. At the end of each period, the nurses received electronic forms to evaluate each of the Rs they worked with and completed the forms during the following 2 weeks.

Data were collected in a database and processed in order to perform quantitative and qualitative analysis.

Statistics

Quantitative analysis

Validity: the ability of the O-RON to discriminate the

Busch, et al.: Assessing the professional performance of residents in cardiology by nurse department

Nombre de residente:	Año de residencia:
Sector:	Fecha:

Su opinión sobre el desempeño del/de la residente es extremadamente valiosa para el programa de residencia. Teniendo esto en cuenta, complete este formulario según lo que observó e interactuó con él/ella durante este tiempo. No tome en cuenta ni evalúe las interacciones previas que tuvo con este/esta residente. Se combinarán varios formularios completados por residente para garantizar el anonimato.

ESCALA DE PUNTUACION (marcar con una X según corresponda):
 1. OBSERVO VARIOS MOTIVOS DE PREOCUPACIÓN DE POCA IMPORTANCIA O 1 MOTIVO DE PREOCUPACION IMPORTANTE
 2. OBSERVO 1-2 MOTIVOS DE PREOCUPACIÓN DE POCA IMPORTANCIA
 3. NO OBSERVO MOTIVOS DE PREOCUPACIÓN
 4. NO EVALUABLE

	1	2	3	4
Los conocimientos médicos básicos son apropiados para su etapa de formación				
Comprende el plan de tratamiento y los problemas vinculados al cuidado del paciente y es capaz de explicárselos a otras/os profesionales de la salud, pacientes y/o familiares.				
Escribe evoluciones diarias útiles (ej., lee y tiene en cuenta otras evoluciones de la historia clínica). Las mismas son fácilmente interpretables y tienen sentido, provee plan de acción.				
Se dirige a los pacientes con un tono de voz apropiado y es accesible a ellos (ej., lenguaje corporal, dispuesto/a dedicarles tiempo, etc.)				
Se dirige al personal de enfermería y otros profesionales de la salud con un tono de voz apropiado y es comprensible (ej., lenguaje corporal, dispuesto/a dedicarles tiempo, etc.)				
Escucha atentamente a los pacientes (presta atención e identifica los problemas)				
Escucha atentamente al personal de enfermería y a otros profesionales de la salud (presta atención e identifica los problemas)				
Acepta apropiadamente la opinión y la experiencia del personal de enfermería				
Interactúa de manera adecuada y efectiva cuando se comunica por teléfono con el personal de enfermería				
Disponibilidad: responde a los mensajes o se encarga de que sean respondidos en caso de que él no pueda; busca activamente comunicarse frente a frente con el personal de enfermería para resolver los problemas del paciente				
Lleva a cabo tareas de cuidado del paciente a tiempo (es decir, realiza aquello que se dijo que se realizaría sin recordatorios constantes).				
Cumple con los protocolos del sector (ej., aislamiento, técnica estéril, etc.)				
Actúa con honestidad e integridad				
Se esfuerza por crear un clima de trabajo positivo y respetuoso (ej., aprende los nombres de sus compañeros de trabajo, dice gracias, etc.)				
Defiende los derechos de los pacientes sin perjudicar ni desacreditar a otros profesionales de la salud				

¿Cuáles son las fortalezas de este residente?

¿Qué podría mejorar?

Figure 1: O-RON form translated and adapted

level of experience of the physicians according to their postgraduate year level will be evaluated. It would be expected that the skills and knowledge of each R will vary according to his/her postgraduate year level. Data are expressed as percentages and frequencies of the qualifications obtained in the different questions. The Chi-square test was used to assess the significance of the differences obtained.

Reliability: internal consistency will be evaluated using Cronbach’s alpha. Reproducibility will be assessed by using a set of statistical procedures (generalizability theory) to determine if the results obtained in the generalizability analysis are reliable. We will analyze which percentage of variance in the observations corresponds to the evaluators, to the residents, and to their postgraduate year level, and we will determine the number of observations per resident needed to achieve a certain level of reliability.

Feasibility: each resident will require a minimum of four observations per evaluation round.

Satisfaction with the tool: was assessed using a survey with a 10-point scale designed by the authors.

Qualitative analysis

There are two questions before the end of the form to be answered directly by the observer. One is intended to highlight the positive and commendable aspects of the R: “What are the strengths of this resident?”; and in the other, the observer can comment: “What could this resident improve?”. Comments will be grouped into categories assembled from the analysis of an initial sample of forms (between 10 to 20% of the total, depending on the number of forms completed) where those with no concerns identified will be used to identify categories of positive aspects, and those with multiple concerns identified will be used to identify categories of areas to improve.

Ethical considerations

The Rs signed an informed consent form to endorse their participation in the study, agreeing to be observed and evaluated during the study period. They were assured that the results of the study would not affect the promotion criteria of the residency program. The study was reviewed and approved by the ethics committee of both institutions (ethical code number: 5298).

Results

During the eight observation rounds, 46 nurses (37 in ICBA and 9 in HA) evaluated 23 residents (15 in ICBA and 8 in HA). A total of 838 assessments were completed with an average of 36 observations per resident during the study period (range: 7–75). All the residents were evaluated at least once in all the observation rounds.

Quantitative analysis

Validity: the 15-item form could significantly discriminate the experience of the Rs according to their postgraduate year level. For each of these 15 items, there was a significant association between the results of the assessment and the postgraduate year level, so the evaluators detected more major concerns among junior residents than among senior residents ($P < 0.005$) [Table 1].

Reliability: the form demonstrated very high internal consistency (Cronbach’s coefficient of 0.978). Table 2 shows the results of the generalizability test proposed which demonstrates that the relationship between Rs (i.e., our object of evaluation) and observers is responsible for 73% of the variance observed. Table 3 shows the results of a decision study with the resulting reliability as a function of the number of forms per resident.

Feasibility: overall, the 23 residents were assessed 838 times throughout the eight observation rounds, resulting in an average of 4.55 assessments per resident per observation round. This value remained stable during the eight rounds (1st: 4.65; 2nd: 4.34; 3rd: 4.47; 4th: 6.17; 5th: 4.56; 6th: 4.08; 7th: 4.36; 8th: 3.91).

Satisfaction: the tool received a positive evaluation from nurses, with scores ranging from 3 to 10 (mean 7.75 ± 1.71) in the different items they were asked about, demonstrating an acceptable level of satisfaction with the tool.

Qualitative analysis

Out of the 838 forms completed, nurses provided comments about positive or commendable aspects

Table 1: Rating response frequency for each form item by postgraduate year level (PGY)

Residents	Many concerns	Few concerns	Response option	
			No concerns	Unable to assess
Item 1				
PGY-2	13 (7.0%)	53 (28.5%)	98 (52.7%)	22 (11.8%)
PGY-3	6 (1.5%)	36 (8.7%)	301 (73.1%)	69 (16.1%)
PGY-4	18 (7.5%)	9 (3.8%)	186 (77.5%)	27 (11.3%)
Item 2				
PGY-2	15 (8.1%)	70 (37.6%)	83 (44.6%)	18 (9.7%)
PGY-3	7 (1.7%)	35 (8.5%)	304 (73.8%)	66 (16.0%)
PGY-4	17 (7.1%)	14 (5.8%)	182 (75.8%)	27 (11.3%)
Item 3				
PGY-2	17 (9.1%)	53 (28.5%)	84 (45.2%)	32 (17.2%)
PGY-3	8 (1.9%)	46 (11.2%)	292 (70.9%)	66 (16.0%)
PGY-4	18 (7.5%)	15 (6.3%)	179 (74.6%)	28 (11.7%)
Item 4				
PGY-2	17 (9.1%)	44 (23.7%)	105 (56.5%)	20 (10.8%)
PGY-3	10 (2.4%)	47 (11.4%)	287 (69.7%)	68 (16.5%)
PGY-4	17 (7.1%)	23 (24.2%)	172 (71.7%)	28 (11.7%)
Item 5				
PGY-2	20 (10.1%)	45 (24.2%)	101 (50.3%)	20 (10.8%)
PGY-3	7 (1.7%)	48 (11.7%)	291 (70.6%)	66 (16.0%)
PGY-4	17 (7.1%)	23 (9.6%)	172 (71.7%)	28 (11.7%)
Item 6				
PGY-2	19 (10.2%)	54 (29.0%)	90 (48.4%)	23 (12.4%)
PGY-3	9 (2.2%)	52 (12.6%)	290 (70.4%)	61 (14.8%)
PGY-4	17 (7.1%)	17 (7.1%)	181 (75.4%)	25 (10.4%)
Item 7				
PGY-2	19 (10.2%)	58 (31.2%)	89 (47.8%)	20 (10.8%)
PGY-3	13 (3.2%)	79 (19.2%)	256 (62.1%)	64 (15.5%)
PGY-4	17 (7.1%)	29 (12.1%)	166 (69.2%)	28 (11.7%)
Item 8				
PGY-2	19 (10.2%)	57 (30.6%)	88 (47.3%)	22 (11.8%)
PGY-3	13 (3.2%)	88 (21.4%)	242 (58.7%)	69 (16.8%)
PGY-4	18 (7.5%)	37 (15.4%)	158 (65.8%)	27 (11.3%)
Item 9				
PGY-2	20 (10.8%)	48 (25.8%)	95 (51.1%)	23 (12.4%)
PGY-3	10 (2.4%)	34 (8.3%)	304 (73.8%)	64 (15.5%)
PGY-4	17 (7.1%)	16 (6.7%)	179 (74.6%)	28 (11.7%)
Item 10				
PGY-2	22 (11.8%)	57 (30.6%)	90 (48.4%)	17 (9.1%)
PGY-3	14 (3.4%)	65 (15.8%)	270 (65.5%)	63 (15.3%)
PGY-4	19 (7.9%)	28 (11.7%)	167 (69.6%)	26 (10.8%)
Item 11				
PGY-2	18 (9.7%)	67 (36.0%)	75 (40.3%)	26 (14.0%)
PGY-3	16 (3.9%)	83 (20.1%)	247 (60.0%)	66 (16.0%)
PGY-4	17 (7.1%)	40 (16.7%)	156 (65.0%)	27 (11.3%)
Item 12				
PGY-2	21 (11.3%)	45 (24.2%)	95 (51.1%)	25 (13.4%)
PGY-3	9 (2.2%)	41 (10.0%)	296 (71.8%)	66 (16.0%)
PGY-4	18 (7.5%)	69 (28.8%)	127 (52.9%)	26 (10.8%)
Item 13				
PGY-2	18 (9.7%)	40 (21.5%)	108 (58.1%)	20 (10.8%)
PGY-3	6 (1.5%)	13 (3.2%)	311 (75.5%)	82 (19.9%)
PGY-4	17 (7.1%)	8 (3.3%)	182 (75.8%)	33 (13.8%)

Contd...

Table 1: Contd...

Residents	Many concerns	Few concerns	Response option	
			No concerns	Unable to assess
Item 14				
PGY-2	19 (10.2%)	42 (22.6%)	98 (52.7%)	27 (14.5%)
PGY-3	8 (1.9%)	34 (8.3%)	300 (72.9%)	70 (17.0%)
PGY-4	18 (7.5%)	15 (6.3%)	175 (72.9%)	32 (13.3%)
Item 15				
PGY-2	17 (9.1%)	37 (19.9%)	100 (53.8%)	32 (17.2%)
PGY-3	8 (1.9%)	21 (5.1%)	305 (74.0%)	78 (18.9%)
PGY-4	16 (6.7%)	15 (6.3%)	172 (71.7%)	37 (15.4%)

Table 2: G study

Facet	Variance	% variance
O	210.43	3.4
O/R	2750.64	73.1
I	23.69	0.6
RI	38.28	0.4
OI/R	774.83	22.5
Total	3797.87	100%

O: observer; R: resident; I: item

715 times (85%) and comments about areas to improve 532 times (63%).

For both positive aspects and areas to improve, responses typically referred to aspects of interpersonal skills (e.g., politeness, empathy, teamwork, respect, friendliness), responsiveness and reliability (e.g., ability to solve problems, self-confidence, organization), communication skills (e.g., active listening, well-disposed), and knowledge [see Table 4 for themes and sample responses]. There were no written comments about the study, the form, or its design.

Discussion

There is a strong need for transforming the way residents are assessed toward the evaluation of professional competence to meet the expected requirements for the practice of the profession.^[2,3] The tools currently used are often designed by those who evaluate Rs. 360-degree assessments may be designed to capture nurses' perspective in the evaluation of Rs,^[11] but there is evidence that this may not be considered when combined with the assessments performed by physicians.^[12,13] In addition, as we have previously mentioned, physicians perform differently when they are assessed by peers, and the way in which physicians assess their peers is different from that of nurses.^[8,14] Moreover, many areas of Rs' training are directly observed by nurses but not by their superiors.^[8]

In addition, current multisource feedback tools often use traditional Likert-style anchors and they tend to

find scores that cluster around the top end of the scale.^[15] Research suggests that WPBA tools using entrustment anchors provide more reliable assessments than traditional anchors,^[16] but there is a need for evidence that describes how and why entrustment anchors work. The goal of the original study was to address concerns by developing a tool that rates the performance of a resident on a hospital ward from the perspective of the nurse, using their language and framework for physician competence.

For all these reasons, we believe that the implementation of a tool designed by and for nurses offers a more realistic and authentic assessment of residents' performance in the workplace setting.^[9]

The 15 items identified by the nurses and the responses obtained in the open-ended questions represent several key areas of residents' performance: interpersonal skills, responsiveness, reliability, communication, and knowledge. These are consistent with previous studies that identified patient and family education, interpersonal communication, and professionalism as areas in which nurses have the opportunity to evaluate residents.^[8,12] Our results indicate that residents meet expectations in these areas in most observations made by nurses.

The validity of this tool became evident because each item could discriminate the level of performance and experience of the residents, as the reasons for concern were less common as the postgraduate year increased.

Most feedback tools have ratings clustered around the top end of the scale where everyone is "at or above expectations," thereby hindering the identification of the resident's performance. In this study, the assessors were willing to identify concerns and write comments about the areas to improve. Probably, it is easier for raters to assess these items from the perspective of whether they have "concerns" than with more traditional rating scales that ask them to rate items as compared to their "expectations." In addition, the specificity of the items to evaluate may have helped raters to identify concerns (ex. tone of voice, active listening, etc., as opposed to a general area like "communication skills"). This particular item should be studied in future experiences.

In this experience, the average number of evaluations per resident and assessments per resident per evaluation round was much higher than in the original study (36 vs. 28 and 4.55 vs. 3.15, respectively), even with fewer nurses and fewer residents than in the original study. In our decision study, we determined that

60 forms per resident are required to achieve reliability (G coefficient >0.7) while at least 30 are required to have an at least acceptable accuracy (0.57). A priori, these numbers seem difficult to achieve under current conditions. The difference in figures from the original experience may be due to differences in the design of the decision study. Dudek *et al.* considered the evaluation rounds, observers, and residents and forms (as nested factors) whereas in our study residents are nested with observers and we did not include the evaluation rounds as a facet in the analysis.

As in the original experience, completing the O-RON was voluntary with no incentives provided. During the study period, the hospital instituted a new electronic health record, which represented a substantial new amount of work for all hospital employees to learn the new system.

Despite these factors, there was excellent participation and a large number of forms were completed. This suggests that the form was easy to complete and that the nurses were interested in providing their perspective with regard to resident performance. Ideally, this means that the nurses would be willing to continue with a similar participation rate outside of the study; however, that will need to be examined in a future study.

One of the limitations of this study is that it was carried out in only two centers of the same specialty. It would

Table 3: Number of forms required to achieve reliability

Number of assessments per resident per round	G coefficient
60	0.72
50	0.69
40	0.64
30	0.57
20	0.47

Table 4: Examples of positive aspects of residents and areas to improve

Positive aspects	Examples	Aspects to improve	Examples
Interpersonal skills	"Helps to solve problems and is approachable." "Shows empathy with patients and family" "Disposed to generate a good work environment."	Interpersonal issues	"Providing better information and improving teamwork" "Listening more to nurses." "Active listening to nurses. Is not an easily approachable professional."
Communication skills	"Great communication with the health care team." "Good disposition and actively listens with nurses" "Communicates the changes made."	Communication issues	"Arrogant attitude while communicating with people." "Providing better information and improving teamwork" "Better and fluid communication with nurses."
Responsiveness and reliability	"Accepts nurse's suggestions, good management of critical situations."	Responsiveness and reliability	"Needs to spend more time in the coronary care unit." "Unsure when dealing with an emergency. Needs to cope with nerves."
Knowledge	"Very knowledgeable and shares knowledge."		

be interesting to evaluate this tool in multiple clinical scenarios. On the other hand, the validity analysis of this tool was limited because the group in charge of the translation and cross-cultural adaptation of the original form decided not to include a final question on whether the nurse wanted the resident on his or her team as a global judgment parameter.

The strength of this study lies in the fact that we were able to replicate an original evaluation strategy in a socially and culturally different environment, obtaining similar results. Even having eliminated a global evaluation question that the original form included.

Conclusion

The WPBA tool O-RON can provide residents with a valuable source of feedback from the eyes of nurses on important aspects of their professional training.

This tool, positively assessed by the raters, significantly discriminates against residents' experience. Its implementation is feasible in our environment, and it is user-friendly, though it requires a considerable number of assessments to achieve high reliability.

Acknowledgments

The researchers would like to express their gratitude to the participants both residents and nurses without whom it would not have been possible to carry out this experience.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Nair BKR, Moonen-van Loon JM, Parvathy M, Jolly BC, van der Vleuten CP. Composite reliability of workplace-based assessment of international medical graduates. *Med J Aust* 2017;207:453.
2. Embo M, Driessen E, Valcke M, van der Vleuten CPM. Integrating learning assessment and supervision in a competency framework for clinical workplace education. *Nurse Educ Today* 2015;35:341–6.
3. Batra R, Batra P, Verma N, Bokariya P, Garg S, Yadav S. Mini clinical evaluation exercise (Mini-CEX): A tool for assessment of residents in department of surgery. *J Educ Health Promot* 2022;11:253.
4. Alves de Lima A, Barrero C, Baratta S, Castillo Costa Y, Bortman G, Carabajales J, *et al.* Validity, reliability, feasibility and satisfaction of the Mini-Clinical Evaluation Exercise (Mini-CEX) for cardiology residency training. *Med Teach* 2007;29:785–90.
5. Jafarpoor H, Hosseini M, Sohrabi M, Mehmanavazan M. The effect of direct observation of procedural skills/mini-clinical evaluation exercise on the satisfaction and clinical skills of nursing students in dialysis. *J Educ Health Promot* 2021;10:74.
6. Williams RG, Klamen DA, McGaghie WC. Cognitive, social and environmental sources of bias in clinical performance ratings. *Teach Learn Med* 2003;15:270–92.
7. LaDonna KA, Hatala R, Lingard L, Voyer S, Watling C. Staging a performance: Learners' perceptions about direct observation during residency. *Med Educ* 2017;51:498–510.
8. Ogunyemi D, Gonzalez G, Fong A, Alexander C, Finke D, Donnon T, *et al.* From the eye of the nurses: 360-degree evaluation of residents. *J Contin Educ Health Prof* 2009;29:105–10.
9. Dudek N, Duffy MC, Wood TJ, Gofton W. The Ottawa Resident Observation Form for nurses (O-RON): Assessment of resident performance through the eyes of the nurses. *J Surg Educ* 2021;78:1666–75.
10. Ortiz-Gutiérrez S, Cruz-Avelar A. Proceso de traducción y adaptación cultural de instrumentos de medición en salud. *Actas Dermosifiliogr* 2018;109:202–6.
11. Donnon T, Al Ansari A, Al Alawi S, Violato C. The reliability, validity, and feasibility of multisource feedback physician assessment: A systematic review. *Acad Med* 2014;89:511–6.
12. Byrd A, Theagwara K, McMahon P, Bolton M, Roy M. Using multisource feedback to assess resident communication skills: Adding a new dimension to milestone data. *Ochsner J* 2020;20:255–60.
13. Roberts JK, Sparks MA, Lehigh RW. Medical student attitudes toward kidney physiology and nephrology: A qualitative study. *Ren Fail* 2016;38:1683–93.
14. Probyn L, Lang C, Tomlinson G, Bandiera G. Multisource feedback and self-assessment of the communicator, collaborator, and professional CanMEDS roles for diagnostic radiology residents. *Can Assoc Radiol J* 2014;65:379–84.
15. Moonen-van Loon JMW, Overeem K, Govaerts MJB, Verhoeven BH, van der Vleuten CPM, Driessen EW. The reliability of multisource feedback in competency-based assessment programs: The effects of multiple occasions and assessor groups. *Acad Med* 2015;90:1093–9.
16. Dudek N, Gofton W, Rekman J, McDougall A. Faculty and resident perspectives on using entrustment anchors for workplace-based assessment. *J Grad Med Educ* 2019;11:287–94.