medicina *intensiva*

ORIGINAL

Acceptance and validity of the methods used to implement a competency based medical education programme in an Intensive Care Department of a teaching referral center[☆]

A. Castellanos-Ortega[a], M.J. Broch[a], M. Barrios[a], M.C. Fuentes-Dura[b], M.D. Sancerni-Beitia[b], C. Vicent[a], R. Gimeno[a], P. Ramírez[a], F. Pérez[a], R. García-Ros[c,*]

[a] *Servicio de Medicina Intensiva, Hospital Universitario y Politécnico la Fe, Valencia, Spain*
[b] *Departamento de Metodología de Ciencias del Comportamiento, Universidad de Valencia, Valencia, Spain*
[c] *Departamento de Psicología Evolutiva y de la Educación, Universidad de Valencia, Valencia, Spain*

Abstract
Objectives: 1. To determine the satisfaction of tutors and residents with a specific methodology used to implement CoBaTrICE. 2. To determine the reliability and validity of the global rating scales designed *ad hoc* to assess the performance of the residents for training purposes.
Design: Prospective cohort study.
Participants: All the residents and tutors of the ICU Department of the Hospital Universitario y Politécnico la Fe de Valencia.
Intervention: CoBaTrICE implementation started in March 2016, it was based on: (1) Training the tutors in feedback techniques; (2) Performing multiple objective and structured work based assessments to achieve the competences of the program; and (3) The use of an electronic portfolio to promote learning reflection and to collect the evidence that learning was taking place.
Methods: The acceptance of CoBaTrICE was explored through a satisfaction survey conducted after 9 months of implementation of the training program. The 15 residents and 5 tutors of the ICU Department were asked about the methodology of the formative assessments, the quality of the feedback, self-learning regulation and the electronic portfolio usefulness. The validity of the global rating scales was assessed through the tests alfa de Cronbach, reliability and generalizability indexes, and intraclass correlation coefficient.

Results: The implementation of CoBaTrICE was satisfactory in all the dimensions studied. The global rating scales used for formative purposes showed reliability and validity.

Conclusions: The methodology used to implement CoBaTrICE was highly valued by tutors and residents. The global rating scales used for formative purposes showed reliability and validity.

© 2019 Elsevier España, S.L.U. and SEMICYUC. All rights reserved.

Análisis de la aceptación y validez de los métodos utilizados para la implementación de un programa de formación basado en competencias en un servicio de Medicina Intensiva de un hospital universitario de referencia

Resumen

Objetivos: (*1*) Determinar la satisfacción de tutores y residentes con la metodología utilizada para la implementación de CoBaTrICE, y (*2*) determinar la validez y la fiabilidad de las escalas de valoración global diseñados *ad hoc* para analizar el desempeño de los residentes con fines formativos.

Diseño: Estudio prospectivo de cohortes.

Participantes: Todos los residentes y tutores del Servicio de Medicina Intensiva del Hospital Universitario y Politécnico La Fe de Valencia.

Intervención: En marzo del 2016 se inició la implementación de CoBaTrICE sustentada en: (*1*) formación de los tutores en técnicas de retroalimentación; (*2*) realización por los residentes de múltiples ejercicios reales de evaluación objetiva y estructurada para adquirir las competencias del programa, y (*3*) uso de un portafolio electrónico para registrar las evidencias del progreso y estimular la reflexión.

Métodos: La satisfacción con CoBaTrICE se evaluó mediante una encuesta realizada tras 9 meses de implementación a los 15 residentes y 5 tutores del servicio. Se preguntó sobre la metodología de las evaluaciones, calidad de la retroalimentación, autorregulación del aprendizaje y utilidad del portafolio. Se determinaron la consistencia interna (alfa de Cronbach), índices de generalizabilidad y fiabilidad interjueces (índice de correlación intraclase) de las escalas de valoración global.

Resultados: La aplicación de CoBaTrICE fue satisfactoria en todas las dimensiones estudiadas. Se constataron la validez y la fiabilidad de las escalas de valoración utilizadas.

Conclusiones: La metodología utilizada para implementar CoBaTrICE fue valorada positivamente por tutores y residentes. Las escalas de valoración global utilizadas en la evaluación formativa demostraron ser válidas, fiables y reproducibles.

© 2019 Elsevier España, S.L.U. y SEMICYUC. Todos los derechos reservados.

Introduction

Medical education is currently undergoing a model change. The new paradigm seeks to be more effective, more integrated within the healthcare system, and focused on the direct application of knowledge in clinical practice[1,2]. This new model based on the gradual acquisition of clearly defined, observable and measurable competencies (competency-based medical education [CBME]) is becoming an international reference for defining the outcomes, methods and organization of current medical education[3,4]. Compared with traditional education, which assumes that mere exposure to clinical experiences based on temporary rotations through different Departments suffices to acquire the necessary professional competencies, CBME proposes more solid principles oriented toward satisfying the quality demanded by current society. Some of these principles are[3]: (*a*) defining the learning outcomes that must be shown by residents at the end of their training period; (*b*) focusing attention upon the development and demonstration of skills, attitudes and knowledge acquired by residents in the course of the training process; (*c*) assuming a resident-centered teaching methodology, with less importance being placed on the time dimension of education; (*d*) prioritizing training evaluation and constructive feedback centered on the performance levels of the residents in the real-life working context; and (*e*) using a broad range of evaluation tools and methods. In Intensive Care Medicine, a group of experts in the year 2004 defined for the first time at Spanish national level the competencies to be acquired by professionals wishing to dedicate themselves to the care of the critically ill[5]. Posteriorly, at European level, a program was developed that contemplated all the previously mentioned principles: the Competency Based Training in Intensive Care Medicine in Europe (CoBaTrICE; www.cobatrice.org) initiative, which includes a bundle of 102 minimum competencies

defining work in this particular specialty, and which was born out of the consensus reached by over 500 clinicians and 1400 patients and relatives. The aim of this European initiative was to harmonize the training programs in Intensive Care Medicine in all the countries in order to ensure quality and facilitate the translational flow of professionals[6−9]. Competency Based Training in Intensive Care Medicine in Europe has been adopted by 15 European countries and is in the process of adoption by a similar number of additional countries[5,10], including Spain, within the current legislative framework[11,12].

Although some medical specialties have already defined their programs based on lists of competencies, the CBME model has not yet become the predominant training method because its application requires organizational and cultural changes, resources and particularly more dedication of teaching time, as well as the training of tutors and staff members in education evaluation and feedback techniques[13−15]. Research in this field is still limited, since the model has been applied in a partial and scantly structured manner[4]. There is no validated standard model for implementing CBME, though the key elements for its effective application are[16,17]: (a) the active participation of residents in their training processes; (b) the supervision, follow-up and frequent evaluation of the acquisition of competencies of the residents on the part of their tutors; and (c) the availability of valid and reliable evaluation tools. In line with these requirements, the present study was carried out with the following objectives: (1) to determine the degree of satisfaction of tutors and residents with the methodology used for implementing CoBaTrICE and (2) to determine the validity and reliability of the global assessment scales used by the tutors to evaluate resident performance in the real-life clinical setting. As a more general objective, the study sought to standardize a method for the implementation of CoBaTrICE based on the satisfaction of residents and tutors, and including reliable and reproducible performance assessment scales.

Methods and participants

The study was approved by the Ethics Committee of La Fe Health Research Institute (*Instituto de Investigación Sanitaria La Fe*, Valencia, Spain), and consent was obtained from all the participants (15 residents and 5 tutors).

Study setting

The study was carried out in the Department of Intensive Care Medicine (DICM) of *Hospital Universitario y Politécnico La Fe* (Valencia, Spain) – a reference hospital of the triple-province Valencian Community. The DICM cares for 2200 adult patients/year with mainly medical disorders. The Department has 32 beds distributed between the Intensive Care Unit (ICU) (24 beds) and the Intermediate Care Unit (8 beds). The staff comprises 15 physicians, and of these, 5 are resident tutors accredited by the Teaching Commission of the hospital. The physician/patient ratio is 1:3−4 during the normal work shift, and two staff physicians and two residents in training are present during the on duty periods. The nurse/patient ratio is 1:2. The Department has three

residents in each of the training program years. The specialty program in turn covers a period of 5 years and is distributed into three stages: stage 1 (S1) (R1−R2): 6 residents; stage 2 (S2) (R3−R4): 6 residents, and stage 3 (S3) (R5): 3 residents.

Intervention

The implementation of CoBaTrICE was started in the DICM in March 2016, based on three actions:

1. Training of the tutors. Training evaluation oriented toward the improvement of learning and performance levels is the cornerstone of the CBME model[18], and its core element is constructive feedback delivered to the resident by the tutor and the staff members. For this reason, all the physicians in the Department received a 12-h clinical feedback course based on simulations and role plays imparted by acknowledged experts in this field[19].

2. Multiple objective and structured training assessment exercises based on direct observation of workplace performance at the patient bedside[20−22]. The assessments involved a clinical evaluation mini-exercise (mini-CEX), the assessment of acute care patients (Acute Care Assessment Tools [ACAT]), and the direct observation of procedural performance (Direct Observation Procedural Skills)[23,24]. A global assessment scale was designed to evaluate performance in the management of clinical cases, with another specific scale for the evaluation procedures and techniques, with the purpose of facilitating objective and structured communication between the tutor and resident during the feedback sessions[24,25]. The items and scores assigned to each item of the assessment scales were selected by a group of 6 tutors in Intensive Care Medicine based on the Delphi method. The final model was generated by an iterative test model and evaluation of the scales, with the participation of 6 tutors, three fifth year residents, and an expert in education psychology. Individualized improvement plans were established after each feedback session.

3. Creation of an electronic portfolio to facilitate reflection and facilitate management of the learning process on the part of the resident and the tutor (E-MIR-INTENSIVE®, IIS La Fe, https://emir-intensive.portres.i3net.is/). The portfolio is the property of the resident and registers all the results of the training evaluations, plans, comments and training and healthcare activities carried out. The portfolio also includes a series of short intervention protocols in response to different disease conditions and for the performance of routine techniques and procedures in Intensive Care Medicine. The protocols were developed by the members of the Department and are based on the most current clinical guides applicable to each case. The protocols help the residents to know exactly what they are expected to do in each situation, and seek to reduce clinical variability. Prior to implementation of the program, training lectures were held for all members of the Department, seeking to facilitate comprehension of the principles of the new teaching model, the structure of CoBaTrICE, and the use of the previously mentioned training evaluation tools[26,27].

Statistical analysis

Objective 1. In order to determine the degree of satisfaction of the tutors and residents with the methods used to introduce CoBaTrICE, in December 2016 and 9 months after implementation of the program, a 22-item questionnaire was administered on an anonymous and voluntary basis to all the tutors ($n = 5$) and residents ($n = 15$) in order to explore the four core dimensions of the CBME model (Table 1): methodology of the training evaluations (MTE), quality of feedback provided by the tutors (QFT), promotion of learning self-regulation (PLS) and usefulness of the electronic portfolio (UEP). A 5-point Likert scale was used to score the answers (1 = ''Strongly disagree'', 5 = ''Strongly agree''). The questionnaire also comprised two open-ended questions to address aspects referred to improvement of the program. Selection of the initial repertoire of items of the questionnaire was based on the review of previous studies on the dimensions defining the CBME model[3,4], as well as on the evaluation of learning environments[28] and the application of portfolios in specialized medical education. Following discussion within the research team regarding relevance, applicability, absence of redundancies and clarity in drafting the items, consensus was reached on the final version of the administered questionnaire.

The basic descriptors per items and per dimensions of the questionnaire were obtained after evidencing the internal consistency of the tool based on the Cronbach alpha test and the bivariate associations between them using Kendall's tau measure. Comparative analyses were made of the responses of the residents and of the tutors based on the nonparametric Kruskal–Wallis test, with *a posteriori* comparisons between groups (2×2) using the Mann–Whitney U-test with Bonferroni correction.

Objective 2. Between April and June 2017, an additional study was made to determine the validity and reliability of the global assessment scales designed on an *ad hoc* basis to measure the resident performance levels. The aim was to determine possible significant differences in the results obtained in the evaluations of the residents dependent upon the observing tutor, self-evaluation (SE) of the resident, and the training stage of the latter. Two residents per training stage from R2 onwards participated (S1 [R2]: 2 residents, S2 [R3–R4]: 2 residents and S3 [R5]: 2 residents). The tutors, in randomized groups of three each, simultaneously evaluated the performance of each participating resident in the management of three clinical cases (acute coronary syndrome, septic shock, and acute respiratory failure). The forms corresponding to the assessment scale and the intervention protocols can be consulted at https://emir-intensive.portres.i3net.is/. The residents self-evaluated their performance with the same model, integrating 15 performance criteria based on a 6-point response scale scored according to the existence during the exercise of major omissions or errors with a clear potential influence upon the patient prognosis (1–3 points), minor omissions or errors with no influence upon the prognosis (4–5 points), and complete and orderly performance (6 points). The external evaluations made by the tutors (ET) and the SE of the residents yielded 96 different assessments (72 ET and 24 SE).

The descriptors of the assessments made by the tutors and residents were obtained. Both ET ($p = 0.30$) and SE by the residents ($p = 0.98$) were seen to be homogeneous among the different clinical cases. Considering the global evaluations, we analyzed the structure of the scale based on principal axis factor analysis; its reliability based on internal consistency analysis; and the generalizability index and inter-rater reliability based on calculation of the intraclass correlation coefficient. Calculation of the ET-SE relationship was made based on the bivariate correlation index. Factorial (3×2) analysis of variance (ANOVA) was used to analyze the effect upon the results of the evaluations of the educational stage of the residents (S1: R2; S2: R3–R4; S3: R5) and the type of evaluation (SE *versus* ET).

Results

Results of objective 1

The 15 residents and 5 tutors answered 100% of the questions (Table 1). The basic descriptors, internal consistency and bivariate relationships among the dimensions of the questionnaire for assessment of the implantation of CoBaTrICE are detailed in Table 2. All the dimensions showed high mean scores (range 3.79–4.28) and adequate internal consistency (range 0.68–0.81), evidencing their relative independence (significant correlations: $T_{MTE-UEP} = 0.34$, $p < 0.05$; $T_{PLS-UEP} = 0.54$, $p < 0.01$). The assessments of the residents and tutors referred to each item of the questionnaire yielded high scores (range of scores among residents: 3.27–4.53; range of scores among tutors: 4.0–5.0). The residents assigned the highest assessment scores to the following items: ''Allows residents and tutors to reflect upon the areas of learning where further improvement is needed'' (MTE), ''The evaluation forms incorporate tools that are useful for correctly addressing the clinical cases'' (MTE), ''Tools are included that are very useful for evaluating training progression of the residents'' (MTE) and ''Better documentation and organization of the evidences of my training are allowed'' (UEP). The tutors in turn assigned the highest assessment scores to the following three items of the UEP: ''It facilitates the documentation of the training experiences and achievements of the residents'', ''It affords a global view of the resident training process, as well as of the strengths and weaknesses'' and ''It enhances awareness of the scope (volume and types of cases addressed) of the training experiences''. In contrast, the residents assigned lower scores to the items: ''The tutors are better able to adjust their behavior to the characteristics of each resident'' (QFT), ''Tutor feedback is richer and adjusted to the performance level of the residents'' (QFT) and ''Autonomous learning is encouraged'' (PLS) – with means (*M*) of 3.27, 3.57 and 3.60, respectively.

The Kruskal–Wallis test (Table 3) showed the assessments of the residents and tutors to be homogeneous in MTE, PLS and UEP, while the tutors assigned higher scores to QFT (4.35 ± 0.5 points *versus* 3.59 ± 0.3; $p < 0.02$). The Mann–Whitney U-test in turn revealed significant differences in QFT between the scores of the tutors and residents referred to the first (S1 < tutors; $p < 0.01$) and second training

Table 1 Basic descriptors and distribution by levels of the responses of the participants referred to the items of the questionnaire.

Items	Frequency responses by level					Descriptors				
	1	2	3	4	5	M	SD	Range	As	Ku
Methodology of the training evaluations										
The forms of the structured evaluations include the key competencies for professional practice	0	0	2	9	9	4.35	0.67	3–5	−0.55	−0.55
It is a good guide for allowing tutors to more objectively assess the performance of residents	0	0	3	11	6	4.15	0.67	3–5	−0.18	−0.55
The evaluation forms incorporate tools that are useful for correctly addressing the clinical cases	0	0	0	13	7	4.35	0.49	4–5	0.68	−1.72
The evaluation forms are clearly worded	0	0	3	9	8	4.25	0.72	3–5	−0.41	−0.83
It allows more complete evaluation than the traditional system	0	0	0	12	8	4.40	0.50	4–5	0.44	−2.01
It allows fairer evaluation than the traditional system	0	0	4	11	5	4.05	0.69	3–5	−0.66	−0.63
Tools are included that are very useful for evaluating training progression of the residents	0	0	1	11	8	4.35	0.59	3–5	−0.21	−0.55
Promotion of learning self-regulation										
It facilitates autonomous learning among the residents	0	2	8	7	3	3.55	0.89	2–5	0.08	−0.53
It helps residents to recognize their strengths/weaknesses in professional practice	0	0	4	12	4	4.00	0.65	3–5	0.00	−0.28
It promotes more active learning	0	0	3	12	5	4.10	0.64	3–5	−0.08	−0.25
It helps residents to reflect upon the objectives and achievements reached in their training process	0	0	3	13	4	4.05	0.61	3–5	−0.01	0.19
It allows residents and tutors to reflect upon the areas of learning where further improvement is needed	0	0	0	10	10	4.50	0.51	4–5	0.00	−2.24
Quality of the feedback provided by tutors										
It allows the feedback sessions with the tutors to be more productive	0	0	4	14	2	3.90	0.55	3–5	−0.08	0.77
It allows communication to be more specific and centered on the needs of the residents	0	0	6	12	2	3.80	0.62	3–5	0.12	−0.21
The tutors are better able to adjust their behavior to the characteristics of each resident	0	2	7	10	1	3.50	0.76	2–5	−0.40	−0.04
Tutor feedback is richer and adjusted to the performance level of the residents	0	0	6	12	2	3.80	0.62	3–5	0.12	−0.21
Usefulness of the electronic portfolio										
The material and technological support of the e-portfolio is very adequate	0	0	5	9	6	4.05	0.76	3–5	−0.09	−1.15
Better documentation and organization of the evidences of training are allowed	0	0	1	9	10	4.45	0.61	3–5	−0.58	−0.46
It facilitates documentation of the training experiences and achievements of the residents	0	0	4	5	11	4.35	0.81	3–5	−0.77	−1.00
Recommendation of the e-portfolio to residents of other hospitals	0	1	5	6	8	4.05	0.95	2–5	−0.52	−0.79
It affords a global view of the resident training process, as well as of the strengths and weaknesses	0	0	4	8	8	4.20	0.77	3–5	−0.37	−1.13
It enhances awareness of the scope (volume and types of cases addressed) of the training experiences	0	0	3	11	6	4.15	0.67	3–5	−0.18	−0.55

As: asymmetry; SD: standard deviation; Ku: kurtosis; *M*: mean; 1: strongly disagree; 2: quite disagree; 3: neither agree nor disagree; 4: quite agree; 5: strongly agree.

Table 2 Basic descriptors (*M* = mean; SD = standard deviation), correlations matrix (Kendall's tau-b) and internal consistency (Cronbach alpha in diagonal) of the subscales of the questionnaire for assessment of the methodology used to implement CoBaTrICE.

Dimension	M	SD	1	2	3	4
1. MTE	4.28	0.36	(0.68)			
2. QFT	3.79	0.50	−0.09	(0.81)		
3. PLS	4.04	0.49	0.21	0.21	(0.78)	
4. UEP	4.21	0.55	0.34*	0.16	0.54**	(0.77)

QFT: quality of feedback provided by the tutors; MTE: methodology of the training evaluations; PLS: promotion of learning self-regulation; UEP: usefulness of the electronic portfolio.
* $p < 0.05$.
** $p < 0.01$.

Table 3 Basic descriptors (mean and standard deviation in parentheses) and significant differences in the 2×2 comparisons between groups in the resident and tutor assessment of the methodology used to implement CoBaTrICE.

	Residents				Tutors	p
	Stage 1	Stage 2	Stage 3	Total		
MTE	4.31 (0.3)	4.19 (0.4)	4.33 (0.2)	4.27 (0.3)	4.29 (0.5)	
PLS	3.97 (0.6)	4.10 (0.5)	3.93 (0.1)	4.01 (0.4)	4.12 (0.5)	
QFT	3.40 (0.4)	3.71 (0.3)	3.67 (0.3)	3.59 (0.3)	4.35 (0.5)	*S1 < tutors *S2 < tutors
UEP	4.44 (0.3)	3.92 (0.7)	4.06 (0.3)	4.16 (0.5)	4.37 (0.6)	

QFT: quality of feedback provided by the tutors; MTE: methodology of the training evaluations; PLS: promotion of learning self-regulation; UEP: usefulness of the electronic portfolio.
* $p < 0.05$.

stages (S2 < tutors; $p < 0.03$). The assessments of the tutors and residents referred to the specific items were homogeneous, with the exception of three corresponding to QFT that were scored higher by the tutors ($p < 0.05$): ''Allows communication to be more specific and centered on the needs of the residents'', ''The tutors are better able to adjust their behavior to the characteristics of each resident'' and ''Tutor feedback is richer and adjusted to the performance level of the residents''.

The responses to the open-ended questions evidenced that the residents and tutors considered the training evaluation methodology compared with the traditional methodology to be characterized by: (*a*) greater objectiveness and use of shared criteria among tutors; (*b*) greater centering of feedback in the performance levels of the residents; and (*c*) better adjustment to their training needs. Among the aspects amenable to improvement, emphasis was placed on the need to: (*a*) standardize the training evaluations timetable; (*b*) improve their compatibility with care overload; (*c*) expand the clinical protocols of the portfolio, and (*d*) improve navigation within the portfolio.

Results of objective 2

Having evidenced data adequacy (Kaiser–Meyer–Olkin [KMO] = 0.87; Bartlett's test of sphericity, chi-squared (66) = 465.1, $p < 0.001$), a principal axis factor analysis was performed. Table 4 reflects its unifactorial structure, presents the descriptors of the items (*M* range 4.55–5.65)

and shows the corresponding factor saturations (range 0.35–0.75). The internal consistency of the scale (Cronbach alpha = 0.88), the discriminating indices of the items (mean total item correlation 0.60) and the generalizability indices ($G_{items} = 0.99$; $G_{subjects} = 0.94$) were adequate. Inter-rater reliability also showed satisfactory values (intraclass correlation coefficient = 0.85, 95% confidence interval [95%CI] 0.78–0.91, $F_{48.672} = 8.1$, $p < 0.001$). The evaluations made by the tutors of the performance levels of the residents in dealing with the clinical cases, and the SE established by the residents themselves, yielded high values (ET: $M = 5.2$, standard deviation [SD] = 0.46; EA: $M = 4.7$, SD = 0.63), evidencing the existence of a significant relationship between both types of evaluation ($r = 0.69$, $p < 0.001$). The results of the ANOVA test revealed the existence of significant principal effects of the type of evaluation upon the assessments derived from them ($F[1.60] = 24.1$, $p < 0.001$, $\eta^2 = 0.29$], with SE exhibiting values significantly lower than for ET. The training stage of the residents also showed significant principal effects ($F[1.60] = 22.1$, $p < 0.001$, $\eta^2 = 0.42$), with the observation of a growing progression of the results of the evaluations between stages ($M_{E1} = 4.5$; $M_{E2} = 5.1$; $M_{E3} = 5.3$) – with significantly higher values in S2–S3 *versus* S1 (differences M: S2–S1 = 0.64, $p < 0.001$; S3–S1 = 0.79, $p < 0.001$).

Discussion

The results of the present study reflect high satisfaction among the residents and tutors with the methodology used

Table 4 Descriptive statistics (mean [*M*], standard deviation [SD], range, asymmetry and kurtosis) and factor saturation of the items of the global assessment scale referred to management of the clinical cases.

Items	Mean	SD	Range	Asymmetry	Kurtosis	Saturation
1. Performs orderly and rapid exploration (ABCD), requests immediate monitoring, identifies severity signs and establishes immediate treatment priorities	4.96	0.71	3–6	−0.31	0.12	0.56
2. Compiles a complete, systematic and problem-oriented case history	5.06	0.63	4–6	−0.04	−0.32	0.75
3. Performs a complete, systematic and problem-oriented physical examination	5.04	0.73	3–6	−0.39	−0.04	0.67
4. Requests and performs the necessary complementary diagnostic tests following a logical order of priorities	5.24	0.63	4–6	−0.24	−0.56	0.72
5. Performs a complete differential diagnosis coherent with the clinical condition of the patient and the provisional results of the complementary tests	5.10	0.62	4–6	−0.07	−0.31	0.73
6. Performs frequent and structured re-evaluation, insisting on the important findings of the previous explorations	5.09	0.58	3–6	−0.74	3.30	0.74
7. Proposes a specific management plan coherent with the clinical condition of the patient and based on the scientific evidence – expressing it clearly and sharing it with the rest of the staff in order to avoid application error	5.09	0.61	4–6	−0.03	−0.17	0.65
8. Correctly manages the steps in time, particularly in situations where the outcome is time-dependent	5.31	0.59	4–6	−0.16	−0.54	0.74
9. Manages resources efficiently (staff, diagnostic tests and treatments)	5.40	0.60	4–6	−0.50	−0.48	0.70
10. Knows his or her limitations, requesting help and advice from other specialists if needed	5.65	0.78	4–6	−1.27	1.29	0.63
11. Prioritizes actions based on their potential risks and benefits	5.22	0.55	4–6	0.09	−0.13	0.35
12. Generates confidence in the team, the patients and their relatives. Effective communication with the team	5.28	0.73	3–6	−0.86	0.70	0.64
13. Shows respect and empathy in communication with the patients and their relatives. Explains the situation and the treatment plan to them in a clear and concise manner, making sure that they have understood the information provided	5.54	0.50	4–6	−0.65	−0.26	0.61
14. Draft a complete, structured and precise clinical report, including the problems and the plan based on priorities	5.21	0.43	4–6	0.77	0.86	0.64
15. Performs an outcome estimate based on the use of prognostic indicators	4.55	0.89	2–6	−0.64	0.48	0.47

to implement CoBaTrICE, as well as adequate reliability and reproducibility of the forms of the global assessment scale used in the training evaluation exercises. The participants positively rated all of the explored dimensions: MTE, PLS, UEP and QFT. The results are particularly significant in relation to the first three dimensions, with strong emphasis on the adequacy of the training evaluations (''They allow a fairer evaluation than the traditional system''), the promotion of autoregulation and reflection of the residents (''Allows residents and tutors to reflect upon the areas of learning where further improvement is needed'') and the usefulness of the portfolio (''It affords a global view of the resident training process, as well as of the strengths and weaknesses''). The training evaluation and consequent feedback provided by the tutor are the key elements for effective implementation of the CBME. The frequent evaluations together with periodic and detailed review of the achievements and pending objectives are crucial for ensuring progression of the residents. In order to take full advantage of the potential of the CBME as a driver of learning, healthcare quality and patient safety, we need effective management of the information and documentation generated during the process, as well as a detailed analysis of the adequacy of the evaluation methods used, with a view to introducing possible modifications in order to secure their increased acceptance among all those involved[29]. In this context, the portfolios are particularly useful tools in the CBME model, since they allow the tutors and residents

improved management of the teaching-learning processes and boost self-awareness, self-guidance and reflection on the part of the residents regarding their performance levels – while also serving as an instrument for quality control of the training process[30,31]. Their use is recommended by the accrediting government institutions of the United Kingdom, the United States and Canada[32,33]. With regard to the quality dimension of feedback, although the assessments of the residents were positive (''It allows the feedback sessions with the tutors to be more productive''), they fell short of the ratings obtained with the rest of the dimensions, and were significantly lower than those of the tutors – particularly in relation to ''The tutors are better able to adjust their behavior to the characteristics of each resident''. This discrepancy should cause us to ask ourselves whether the clinical feedback provided by the tutors was adequate and sufficient, and should lead us to consider the need to maintain it over time or complement it with other actions to promote the analysis and discussion of good practices. Effective feedback requires us not only to focus on performance levels but also to effectively respond to the cognitive, emotional and behavioral dimensions implicated in the activity of the residents. It is necessary to create a stimulating environment in which the professional feels comfortable on reflecting in depth upon the mental model behind his or her actions[34]. The acquisition of these skills can be facilitated through the training of tutors and staff members in coaching techniques, for in addition to serving as instructors, these professionals must act as facilitators and guides to help residents discover their needs, set goals and design realistic strategies for reaching them[35].

A crucial aspect in order for evaluation to be effective is the reliability, accuracy and reproducibility of the measurement instruments used[16]. In this study we chose the global assessment scale model, since it offers greater objectiveness than checklists, which are less costly and easier to use but are also less discriminating[36]. The forms of the global assessment scale created *ad hoc* were seen to afford adequate psychometric features, a unifactorial structure, strong internal consistency, inter-rater reliability in application and generalizability of the results. It can be concluded that they were reliable and valid for assessing the performance level of the residents through direct observation of their actions in real-life settings at the patient bedside, and for providing the pertinent and consequent constructive feedback. We also confirmed the discriminating validity of the assessment scales between training stages, evidencing that the residents in the more advanced stages achieved significantly higher scores in the management of the clinical cases. In this way, although the CBME places greater importance on the progression of the acquisition of competencies than on the time dimension of medical education, significant effects were observed regarding the training stages of the program, and which explain 42% of the variance of the results of the evaluations. The study also evidenced average to high correlation between ET and SE of the residents, evidencing that the assessments of the former are higher than the latter. These results suggest that the CBME model can help residents to make more realistic judgments of their own performance levels[37].

The present study has a number of limitations, including particularly its observational design, the limited number of participants, and the fact that it was carried out in a single institution. The results obtained therefore might not be extrapolatable to other Departments, since a range of factors may influence effective implementation of the CBME[38]. In relation to this latter aspect, it also would be interesting to know the effect of the intervention upon the quality of communication and feedback between the residents and those staff members that are not tutors; this possible association has not yet been explored, due to the gradual incorporation of the latter to the program. Lastly, in order to establish whether CoBaTrICE affords improved competency levels among residents in training than the traditional model based on clinical rotations, a cluster-based randomized trial is being carried out with the participation of 14 Departments of Intensive Care Medicine belonging to 14 Spanish reference hospitals with three residents per year[39]. Despite its limitations, this study provides a model for the implementation of a CBME program that was well accepted by the tutors and residents, and in which use was made of measurement instruments of confirmed validity and reliability[40]. To the best of our knowledge, no publications have been made involving the analysis of acceptance on the part of tutors and residents of a specific methodology for implementing CoBaTrICE, and the validity of a concrete global assessment scale model for training evaluation in this specific setting has not been established to date.

Conclusions

The methodology used for the implementation of CoBaTrICE was positively rated by both the tutors and residents. The latter demanded increased quality of the feedback provided by the tutors. The forms of the global assessment scale used in the training evaluation exercises were seen to be valid and reliable in assessing the performance level of the residents through direct observation of their actions in the real-life setting at the patient bedside.

Authorship

Study conception (C), literature review (LR), questionnaires (QU), study development and data acquisition (D), statistical analyses (S), drafting of the manuscript (Draft), critical review (CR) and final draft (FD).

- Alvaro Castellanos: C, LR, D, Draft, FD.
- Maria Jesus Broch: C, D, CR.
- Marcos Barrios: D, CR.
- Dolores Sancerni: QU, S, CR.
- Maria Castillo: QU, S, CR.
- Carlos Vicent: D, CR.
- Ricardo Gimeno: D, CR.
- Paula Ramirez: D, CR.
- Francisca Perez: D, CR.
- Rafael Garcia-Ros: C, LR, QU, D, S, Draft, FD.

Financial support

Research Institute (*Instituto de Investigación Sanitaria La Fe*, Valencia, Spain) (10-COBATRICE-2015-A, Principal investigator: Alvaro Castellanos-Ortega), and by the program for the promotion of scientific research, technological development and innovation of the Valencian Community, of the *Consellería de Educación, Investigación, Cultura y Deporte* of the *Generalitat Valenciana* (GV-AICO-2018-126, Principal investigator: Rafael Garcia-Ros).

Conflicts of interest

Alvaro Castellanos-Ortega is the creator of the e-MIR-Intensive electronic portfolio and holds participation in the benefits of the application.

M.J. Broch-Porcar, M. Barrios de Pedro, M.C. Fuentes, D. Sancerni, C. Vicent-Perales, R. Gimeno-Costa, P. Ramímez-Galleymore, F. Pérez-Esteban and R. García-Ros declare that they have no conflicts of interest in relation to the present article.

References

1. Gruppen LD, Mangrulkar RS, Kolars JC. The promise of competency-based education in the health professions for improving global health. Hum Resour Health. 2012;10:43, http://dx.doi.org/10.1186/1478-4491-10-43.

2. Castellanos-Ortega A, Rothen HU, Franco N, Rayo LA, Martín-Loeches I, Ramírez P, et al. Training in Intensive Care Medicine. A challenge within reach. Med Intensiva. 2014;38:305–10, http://dx.doi.org/10.1016/j.medin.2013.12.011.

3. Frank JR, Danoff D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. Med Teach. 2007;29:642–7, http://dx.doi.org/10.1080/01421590701746983.

4. Gruppen LD, Burkhardt JC, Fitzgerald JT, Funnell M, Haftel HM, Lypson ML, et al. Competency-based education: programme design and challenges to implementation. Med Educ. 2016;50:532–9, http://dx.doi.org/10.1111/medu.12977.

5. Roca J, Pérez JM, Colmenero M, Muñoz H, Alarcón L, Vázquez G. Professional competence for the care of critical patients: beyond specialities. Med Intensiva. 2007;31:473–84, http://dx.doi.org/10.1016/S0210-5691(07)74854-4.

6. Bion JF, Barrett H. Development of core competencies for an international training programme in intensive care medicine. Intensive Care Med. 2006;32:1371–83, http://dx.doi.org/10.1007/s00134-006-0215-5.

7. CoBaTrICE Collaboration. International standards for programmes of training in intensive care medicine in Europe. Intensive Care Med. 2011;37:385–93, http://dx.doi.org/10.1007/s00134-010-2096-x.

8. CoBaTrICE Collaboration. The views of patients and relatives of what makes a good intensivist: a European survey. Intensive Care Med. 2007;33:1913–20, http://dx.doi.org/10.1007/s00134-007-0799-4.

9. CoBaTrICE Collaboration. The educational environment for training in intensive care medicine: structures, processes, outcomes and challenges in the European region. Intensive Care Med. 2009;35:1575–83, http://dx.doi.org/10.1007/s00134-009-1514-4.

10. Rhodes A, Chiche JD, Moreno R. Improving the quality of training programs in intensive care: a view from the ESICM. Intensive Care Med. 2011;37:377–9, http://dx.doi.org/10.1007/s00134-010-2120-1.

11. Martin-Delgado MC, Gordo-Vidal F. La calidad y la seguridad de la medicina intensiva en España: algo más que palabras. Med Intensiva. 2011;35:201–5, http://dx.doi.org/10.1016/j.medin.2011.02.003.

12. Real Decreto 183/2008, de 8 de febrero, por el que se determinan y clasifican las especialidades de Ciencias de la Salud y se desarrollan determinados aspectos de la formación sanitaria especializada. BOE-A-2008-3176.

13. Frank JR, Snell LS, Cate OT, Holmboe ES, Carraccio C, Swing SR, et al. Competency-based medical education: theory to practice. Med Teach. 2010;32:638–45, http://dx.doi.org/10.3109/0142159X.2010.905011.

14. Nousiainen MT, Caverzagie KJ, Ferguson PC, Frank JR. Implementing competency-based medical education: what changes in curricular structure processes are needed? Med Teach. 2017;39:594–8, http://dx.doi.org/10.1080/0142159X.2017 0771315.

15. Caverzagie KJ, Nousiainen MT, Ferguson PC, Ten Cate O, Ross S, Harris KA, et al. Overarching challenges to the implementation of competency-based medical education. Med Teach. 2017;39:588–93, http://dx.doi.org/10.1080/0142159X.2017.0751315.

16. Iobst WF, Sherbino J, Cate OT, Richardson DL, Dath D, Swing SR, et al. Competency-based medical education in postgraduate medical education. Med Teach. 2010;32:651–6, http://dx.doi.org/10.3109/0142159X.2010.095007.

17. Voorhees RA, Bedard-Voorhees A. Principles for competency-based education. In: Reigeluth Ch, Beatty BJ, Myers RD, editors. Instructional-design theories and models. Volume IV: The learner-centered paradigm of education. New York: Routledge; 2017. p. 33–63.

18. Holmboe ES, Sherbino J, Long DM, Swing SR, Frank JR. The role of assessment in competency-based medical education. Med Teach. 2010;32:676–82, http://dx.doi.org/10.3109/0142159X. 2010. 045007.

19. Maestre JM, Rudolph JW. Theories and styles of debriefing: the good judgment method as a tool for formative assessment in healthcare. Rev Esp Cardiol (Engl Ed). 2015;68:282–5, http://dx.doi.org/10.1016/j.rec.05.0182014.

20. Holmboe ES, Hawkins RE, Huot SJ. Effects of training in direct observation of medical residents' clinical competence: a randomized trial. Ann Intern Med. 2004;140:874–81, http://dx.doi.org/10.7326/0003-4819-140-11-200406010-80000.

21. Epstein RM. Assessment in medical education. N Engl J Med. 2007;356:387–96, http://dx.doi.org/10.1056/NEJMra054784.

22. Kogan JR, Holmboe ES, Hauer KE. Tools for direct observation and assessment of clinical skills of medical trainees: a systematic review. JAMA. 2009;302:1316–26, http://dx.doi.org/10.1001/jama.2009.1365.

23. Morán-Barrios J. La evaluación del desempeño o de las competencias en la práctica clínica. 1.ª parte: principios y métodos, ventajas y desventajas. Educ Med. 2016;17:130–9, http://dx.doi.org/10.1016/j.edumed.2016.07.001.

24. Morán-Barrios J. La evaluación del desempeño o de las competencias en la práctica clínica. 2.ª parte: tipos de formularios, diseño, errores en su uso, principios y planificación de la evaluación. Educ Med. 2017;18:2–12, http://dx.doi.org/10.1016/j.edumed.2016.09.003.

25. Jonsson A, Panadero E. The use and design of rubrics to support assessment for learning. In: Carless D, Bridges SM, Chan CKY, Glofcheski R, editors. Scaling up assessment for learning in higher education. Singapore: Springer; 2017. p. 99–111.

26. Dath D, Iobst W. The importance of faculty development in the transition to competency-based medical education. Med Teach. 2010;32:683–6, http://dx.doi.org/10.3109/0142159X. 2010.500710.

27. Ferguson PC, Caverzagie KJ, Nousiainen MT, Snell L, ICBME Collaborators. Changing the culture of medical training: an important step toward the implementation of competency-based

medical education. Med Teach. 2017;39:599–602, http://dx.doi.org/10.1080/0142159X. 2017.1315079.

28. Schönrock-Adema J, Visscher M, Raat AJ, Brand PL. Development and validation of the Scan of Postgraduate Educational Environment Domains (SPEED): a brief instrument to assess the educational environment in postgraduate medical education. PLOS ONE. 2015;10:e0137872, http://dx.doi.org/10.1371/journal.pone.0137872.

29. Lockyer J, Carraccio C, Chan MK, Hart D, Smee S, Touchie C, et al. Core principles of assessment in competency-based medical education. Med Teach. 2017;39:609–16, http://dx.doi.org/10.1080/0142159X. 2017.1315082.

30. Tochel C, Haig A, Hesketh A, Cadzow A, Beggs K, Colthart I, et al. The effectiveness of portfolios for post-graduate assessment and education: BEME Guide No. 12. Med Teach. 2009;31:299–318, http://dx.doi.org/10.1080/01421590902883056.

31. Heeneman S, Driessen EW. The use of a portfolio in postgraduate medical education – reflect, assess and account, one for each or all in one? GMS J Med Educ. 2017;34:Doc57, http://dx.doi.org/10.3205/zma001134.

32. Donato AA, George DL. A blueprint for implementation of a structured portfolio in an internal medicine residency. Acad Med. 2012;87:185–91, http://dx.doi.org/10.1097/ACM.0b013e340a21823.

33. Hall P, Byszewski A, Sutherland S, Stodel E. Developing a sustainable electronic portfolio (ePortfolio) program that fosters reflective practice and incorporates CanMEDS competencies into the undergraduate medical curriculum. Acad Med. 2012;87:744–51, http://dx.doi.org/10.1097/ACM.0b013e318253dacd.

34. Maestre JM, Szyld D, del Moral I, Ortiz G, Rudolph JW. The making of expert clinicians: reflective practice. Rev Clin Esp (Barc). 2014;214:216–20, http://dx.doi.org/10.1016/j.rce.2013.12.001.

35. Orr CJ, Sonnadara RR. Coaching by design: exploring a new approach to faculty development in a competency-based medical education curriculum. Adv Med Educ Pract. 2019;10:229–44, http://dx.doi.org/10.2147/AMEP.S191470.

36. Ilgen JS, Ma IW, Hatala R, Cook DA. A systematic review of validity evidence for checklists versus global rating scales in simulation-based assessment. Med Educ. 2015;49:161–73, http://dx.doi.org/10.1111/medu.12621.

37. Davis DA, Mazmanian PE, Fordis M, van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. JAMA. 2006;296:1094–102, http://dx.doi.org/10.1001/jama.296.9.1094.

38. Lörwald AC, Lahner FM, Mooser B, Perrig M, Widmer MK, Greif R, et al. Influences on the implementation of Mini-CEX and DOPS for postgraduate medical trainees' learning: a grounded theory study. Med Teach. 2019;41:448–56, http://dx.doi.org/10.1080/0142159X. 2018.1497784.

39. IP: Álvaro Castellanos-Ortega. Available from: https://www.esicm.org/research/trials/trials-group-2/cobalidation/ [accessed 06.01.20].

40. Gruppen L, Frank JR, Lockyer J, Ross S, Bould MD, Harris P, et al. Toward a research agenda for competency-based medical education. Med Teach. 2017;39:623–30, http://dx.doi.org/10.1080/0142159X. 2017.0651315.