



## ORIGINAL

## Concordancia competencial entre el Examen Clínico Objetivo Estructurado y la prueba de casos clínicos computarizados de Medicina: ¿estamos evaluando lo mismo?



Emilio Cervera-Barba<sup>a,g,\*</sup>, José Manuel Blanco-Canseco<sup>b,g</sup>, Pilar Fidalgo-Montero<sup>c,g</sup>, M. Isabel Gámez-Cabero<sup>b,g</sup>, Francisco Hijano-Bandera<sup>d,g</sup>, M. Soledad Holgado-Catalán<sup>d,g</sup>, Irene Salinas-Gabiña<sup>c,g</sup>, Fernando Neria-Serrano<sup>f</sup>, Raúl Castañeda-Vozmediano<sup>f</sup> y Sophia Denizon-Arranz<sup>e,g</sup>

<sup>a</sup> Prácticas y Centros Sanitarios, Facultad de Medicina, Universidad Francisco de Vitoria, Madrid, España

<sup>b</sup> Centro de Salud Valle de la Oliva, Madrid, España

<sup>c</sup> Hospital Universitario del Henares, Madrid, España

<sup>d</sup> Centro de Salud Monterrozas, Madrid, España

<sup>e</sup> Centro de Simulación Clínica Avanzada, Facultad de Medicina, Universidad Francisco de Vitoria, Madrid, España

<sup>f</sup> Unidad de Apoyo a la Investigación, Facultad de Medicina, Universidad Francisco de Vitoria, Madrid, España

<sup>g</sup> Comité de Prueba ECOE, Facultad de Medicina, Universidad Francisco de Vitoria, Madrid, España

Recibido el 12 de abril de 2023; aceptado el 8 de junio de 2023

Disponible en Internet el 21 de julio de 2023

### PALABRAS CLAVE

Examen Clínico  
Objetivo Estructurado;  
Prueba de casos  
clínicos  
computarizados;  
Simulación de casos  
clínicos  
computarizados;  
Pacientes virtuales;  
Estudiantes de  
Medicina;  
Evaluación de la  
competencia

### Resumen

**Introducción:** en el curso 2020-21 las facultades de Medicina españolas realizaron una evaluación con casos-clínicos computarizados (denominada ECOE-CCS) y algunas, además, una ECOE convencional con simuladores y pacientes simulados (EEOE-CS). El Comité ECOE interfacultativo asumió como iguales las competencias evaluadas con ambas pruebas. El objetivo del estudio es comprobar su similitud valorando la concordancia de resultados de alumnos en las competencias evaluadas con las 2 pruebas.

**Métodos:** ciento ocho alumnos participaron en la ECOE-CCS de 10 casos clínicos y una ECOE-CS de 12 estaciones. Hubo 2 versiones de ECOE-CS para disminuir el efecto copiado entre alumnos. La ECOE-CCS tuvo 303 ítems de evaluación, y las ECOE-CS 320 y 321 ítems respectivamente. Los 5 dominios competenciales evaluados con ambas pruebas fueron: anamnesis, exploración clínica, juicio clínico, prevención y profesionalismo. La consistencia interna de ambas ECOE se midió con el alfa de Cronbach. La correlación entre dominios competenciales se testó con el coeficiente de correlación de Pearson (R).

\* Autor para correspondencia.

Correo electrónico: [emilio.cervera@ufv.es](mailto:emilio.cervera@ufv.es) (E. Cervera-Barba).

**Resultados:** los alumnos obtuvieron medias de calificaciones globales similares en ambas pruebas (ECO-ECS: 69,54%; ECO-ECCS: 71,35%). Hubo diferencias significativas en 4 de 5 dominios competenciales evaluados por ambas pruebas. La consistencia interna respecto a sus dominios competenciales fue: ECO-ECS-1  $\alpha=0,62$ ; ECO-ECS-2  $\alpha=0,54$  y ECO-ECCS  $\alpha=0,38$ . No se halló correlación entre calificaciones globales de alumnos en las 2 pruebas ( $R=0,184$ ;  $p=0,056$ ), ni entre los dominios competenciales similares evaluados con ambas ( $R=-0,2$  a  $0,37$ ).

**Conclusión:** los resultados no avalan la similitud de competencias evaluadas con ambas pruebas. Serán necesarios otros estudios para definir las competencias evaluadas con la ECO-ECCS.

© 2023 The Authors. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## KEYWORDS

Objective Structured Clinical Examination; Computer-based case test; Computer-based case simulation; Virtual patients; Medicine students; Competence assessment

## Competence concordance between the Objective Structured Clinical Examination and the Computerized Case-clinical Test in medicine: are we assessing the same thing?

### Abstract

**Background:** In the 2020-21 academic year, Spanish medical schools carried out an assessment with computerised clinical cases (called CCS-OSCE) and some, in addition, a conventional OSCE with simulators and simulated patients (CS-OSCE). The inter-faculty OSCE Committee assumed the competencies assessed with both tests to be equal. The aim of the study is to check their similarity by assessing the concordance of students' results in the competencies assessed with the two tests.

**Methods:** 108 students participated in the CCS-OSCE of 10 clinical cases and a CS-OSCE of 12 stations. There were two versions of the CS-OSCE to reduce the copying effect among students. The CCS-OSCE had 303 items, and the CS-OSCE had 320 and 321 items. The five competency domains assessed with both tests were: anamnesis, clinical examination, clinical judgement, prevention and professionalism. The internal consistency of both OSCE was measured with Cronbach's alpha. Correlation between competency domains was tested with Pearson's correlation coefficient (R).

**Results:** Students obtained similar overall mean scores in both tests (ECO-ECS: 69.54%; ECO-ECCS: 71.35%). There were significant differences in 4 of 5 competency domains assessed by both tests. The internal consistency with respect to their competence domains was: ECO-ECS-1  $\alpha=0.62$ ; ECO-ECS-2  $\alpha=0.54$  and ECO-ECCS  $\alpha=0.38$ . No correlation was found between students' overall scores in the two tests ( $R=0.184$ ;  $p=0.056$ ), nor between the similar competency domains assessed with both tests ( $R=-0.2$  to  $0.37$ ).

**Conclusion:** The results do not support the similarity of competences assessed with both tests. Further studies will be necessary in order to define the competences assessed with the CCS-OSCE.

© 2023 The Authors. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Introducción

La ECOE (Examen Clínico Objetivo Estructurado) fue propuesta por Harden en 1975 y es una metodología de evaluación de la competencia bien establecida y consolidada en su estructura y organización<sup>1,2</sup>. En ella los participantes se enfrentan a situaciones de la práctica clínica en un entorno de simulación, en similares condiciones de dificultad y una rúbrica de evaluación objetiva y uniforme. La validez y fiabilidad de las ECOE han sido estudiadas y dependen de algunas características intrínsecas de la prueba<sup>3-5</sup>. La Conferencia Nacional de Decanos de Facultades de Medicina de España

(CNDFME) estableció la obligatoriedad de la ECOE para evaluar la competencia final de los egresados<sup>6</sup>.

Durante la pandemia COVID-19 las facultades de Medicina tuvieron que implementar una prueba sustitutiva de la ECOE en el curso 2019-20 y crearon un examen *online* de 10 casos clínicos con preguntas a las que se les asignaron, a criterio del Comité ECOE interfacultativo de la CNDFME (Comité ECOE-CND), los dominios competenciales de la ECOE convencional.

En el curso 2020-21, la CNDFME acordó la realización de una prueba ECOE que incluyera una ECOE convencional presencial y una prueba común *online* similar a la desarrollada

en 2019-20. Por aproximación al sistema norteamericano, e inmersos en el actual proceso de acreditación internacional de las facultades de Medicina españolas, a la ECOE convencional, similar al Step-2 CS (*Clinical Skills*), se la denominó ECOE-CS (Casos Simulados), y a la prueba *online* se la denominó ECOE-CCS (Casos-clínicos Computarizados Simulados) por pretender asemejarla al Step-3 CCS (*Computerized-based Cases Simulation*)<sup>9</sup>. El Comité-CND, integrado por los responsables ECOE de todas las facultades, creó los casos clínicos de la ECOE-CCS con el objetivo de evaluar 5 dominios competenciales de la ECOE-CS: anamnesis, exploración clínica, juicio clínico, prevención y promoción de la salud, y profesionalismo (valores profesionales y aspectos ético-legales). Dichos dominios se asignaron a los ítems de la ECOE-CCS por consenso del Comité ECOE-CND, asumiendo que ambas pruebas podían evaluar las mismas competencias. Veintiséis facultades participamos en la ECOE-CCS y 14 organizamos adicionalmente una ECOE-CS.

Los objetivos de este trabajo son: 1) describir los resultados de los alumnos de 6° curso de nuestra facultad en ambas pruebas y 2) analizar la correlación de resultados de alumnos entre los mismos dominios competenciales evaluados con las 2 pruebas para analizar su similitud.

## Material y métodos

La población del estudio fueron los alumnos de 6° curso de Medicina. La ECOE-CCS se realizó el 7 de mayo de 2021. Dispusieron de 130 minutos y constaba de 10 casos clínicos con 84 preguntas y 303 ítems evaluables. Los alumnos la realizaron vigilados en aula, conectados desde un iPad a través de la plataforma Moodle. Los tipos de preguntas de la prueba fueron: con respuestas de opción múltiple (48%), con respuestas anidadas o incrustadas tipo Cloze (31%), verdadero/falso (6%), emparejamiento de opciones (5%), cuestiones para arrastrar y soltar sobre texto (4%) o imagen (4%), y selección de palabras perdidas (2%). La semana previa los alumnos se entrenaron con un simulacro de prueba similar. La distribución de ítems y puntos por casos y competencias se muestran en la [tabla 1](#).

Al día siguiente los alumnos realizaron la ECOE-CS en el Centro de Simulación Clínica Avanzada de la universidad. Constó de 12 estaciones y la realizaron en 120 minutos. Los alumnos fueron asignados aleatoriamente en 5 turnos de 2 ruedas simultáneas. En cada rueda participaron 12 alumnos, con un total de 9 ruedas. Para paliar el efecto copiado entre alumnos de distintos turnos, se crearon 2 versiones ECOE-CS donde la mitad de las estaciones introducían variaciones en los casos clínicos sin alterar el peso de las competencias evaluadas. En la organización se garantizó la no a esarroll ECOE-CS-1 se desarrolló en ó en No hubo comunicación de alumnos entre turnos sucesivos de ECOE. La ECOE-CS-1 se implementó en los turnos 1°, 2° y 4° (total 6 ruedas), y la ECOE-CS-2 en el 3° y 5° turnos (total 3 ruedas). La ECOE-CS-1 se d Ambas tuvieron 7 pacientes simulados, 2 informes clínicos y 3 simuladores. La ECOE-CS-1 tuvo 320 ítems y 903 puntos; la ECOE-CS-2, 321 ítems y 906 puntos. Las tablas de especificaciones se muestran en la [tabla 2](#). La identidad de los alumnos se ocultó a pacientes simulados y evaluadores mediante un código de letras y números que portaban en la bata. Los informes clínicos se corrigieron el mismo día de la

**Tabla 1** Tabla de especificaciones de la ECOE-CCS

Caso	Especialidad	Preguntas	Anamnesis		Exploración clínica		Juicio clínico		Prev. y promoción		Profesionalismo		Totales	
			Ptos.	Ítems	Ptos.	Ítems	Ptos.	Ítems	Ptos.	Ítems	Ptos.	Ítems	Ptos.	Ítems
1	Reumatología	6	40	8	35	8	60	18	15	1	20	4	100	26
2	Trauma	5	30	5	8	8	60	30	20	7	20	4	100	18
3	Pediatría	8	20	4	8	1	59	12	24	5	24	5	100	41
4	Atención primaria	11	9	9	8	4	30	7	20	13	15	3	100	27
5	Ginecología	7	20	7	15	17	51	7	20	7	20	3	100	34
6	Atención primaria	12	10	5	39	17	80	16	15	3	15	3	100	29
7	Cardiología	7	20	5	20	17	63	23	15	3	15	3	100	21
8	Neumología	8	22	17	6	31	76	31	12	3	12	3	100	43
9	Urgencias	13	12	6	30	4	30	6	106	32	20	9	100	40
10	Psiquiatría	7	20	5	127	34	509	150	106	32	55	16	1000	303
Totales		84	203	71	127	11,2%	509	49,5%	106	10,6%	55	5,5%	1000	100%
% sobre total de la prueba			20,3%	23,4%	12,7%	11,2%	50,9%	49,5%	10,6%	10,6%	5,5%	5,3%	100%	100%
N° casos clínicos que evalúan el dominio competencial			10		5		9		6		3			

Ptos: puntos.

**Tabla 2** Tabla de especificaciones de las ECOE-CS-1 y 2. Se ofrecen los datos de ambas ECOE cuando son diferentes

Estación	Especialidad	Anamnesis		Exploración clínica		Comunicación		Habilidades técnicas		Juicio clínico		Prev. y promoción		Relaciones interprof.		Profesionalismo		Totales				
		Ptos.	Ítems	Ptos.	Ítems	Ptos.	Ítems	Ptos.	Ítems	Ptos.	Ítems	Ptos.	Ítems	Ptos.	Ítems	Ptos.	Ítems	Ptos.	Ítems			
1	Pediatría - C	21	9	14	7	20	5	14	4	6	3							75	28			
2	Atención primaria - C					20	5											36	6	56	11	
3	Neumología - E	26	13	8	4	20	5	16	4											70	26	
4	Reumatología - E			8	2			9	3	36	12	4	1					15	5	72	23	
5	Atención primaria - E	28	12			20	5	14	3			20	4							82	24	
6	Oncología - E			49 / 52	20			15	4											64 / 67	24	
7	Atención primaria - C	44	23			20	5	24	9											88	37	
8	Cardiología - C	43	15	10	2	20	5	15	3											88	25	
9	Informe cardiología - C							9	2			26	8	62	23					97	33	
10	Urgencias - C	20	5	20	4			6	1	26	13									82	28	
11	Digestivo - E	24	16	20	10	20	5	12	4			2	2							78	37	
12	Informe digestivo - E													51	24 / 25					51	24 / 25	
	Total por dominio competencial	206	93	129 / 132	49	140	35	134	37	62	25	58	18	113	47 / 48	61	16	903 / 906	320 / 321			
	% sobre total de la prueba	22,8 / 22,7%	29,1 / 29,0%	14,3 / 14,6%	15,3 / 15,3%	15,5 / 15,4%	10,9 / 10,9%	14,8 / 14,8%	11,5 / 11,5%	7,8 / 7,8%	6,9 / 6,8%	6,4 / 6,4%	5,6 / 5,6%	12,5 / 12,5%	14,7 / 15,0%	6,8 / 6,7%	5,0 / 5,0%	100% / 100%	100% / 100%			
	N° estaciones que evalúan el dominio competencial	7	7	7	6	6	3	10	2	3	5	2	3	2	3	3	3	3	3	3	3	3

C: estación común. E: estación específica para cada ECOE 1 y 2. Ptos: puntos.

prueba. Para los registros de evaluación se utilizó el programa GEXECO® soportado en tablet con sistema operativo Android.

Las variables recogidas de cada alumno fueron: A) de la ECOE-CCS: puntuación en cada pregunta, cada caso clínico y cada dominio competencial, y calificación final; B) de la ECOE-CS: versión 1 o 2, rueda y turno, puntuación en cada ítem, cada estación, cada dominio competencial y calificación final.

Las calificaciones finales y en los dominios competenciales se muestran como media  $\pm$  desviación estándar. Las diferencias entre ECOE-CCS, ECOE-CS-1 y ECOE-CS-2, y entre estaciones o casos, se comprobaron mediante el estadístico T de Student. Para comprobar la correlación entre las 2 pruebas y entre casos o estaciones dentro de cada prueba, se calculó el coeficiente de correlación de Pearson (R). La fiabilidad (consistencia interna) de las pruebas respecto a sus respectivos dominios competenciales se calculó con el alfa de Cronbach ( $\alpha$ ).

Para el análisis estadístico se empleó el *software* R v4.1. Se consideró que existían diferencias significativas cuando el p valor fue igual o menor a 0,05.

El protocolo del estudio fue aprobado por el Comité de Ética de la Investigación de la Universidad Francisco de Vitoria.

## Resultados

Realizaron ambas pruebas el 100% de los alumnos (108; 31 hombres, 28,7%, y 77 mujeres, 71,3%). La ECOE-CS-1 fue administrada a 72 alumnos (66,7%) y la ECOE-2 a 36 (33,3%).

Los resultados globales y por dominios competenciales se muestran en la [tabla 3](#). Los alumnos obtuvieron mejores puntuaciones en la ECOE-CCS, tanto en calificación final como en todos los dominios, menos anamnesis.

Las diferencias de puntuaciones entre las ECOE-CS-1 y 2 solo fueron estadísticamente significativas en la calificación final y la mitad de los dominios competenciales evaluados. Analizados los resultados en las estaciones de las ECOE-CS, solo hubo diferencias significativas en 2 de las estaciones comunes y 2 de las específicas, y se hallaron mejores resultados en los turnos de tarde en 5 de las estaciones de la ECOE-CS-1 y 4 de la ECOE-CS-2. Los resultados significativamente mejores en los turnos de tarde en ambas ECOE-CS fueron los de juicio clínico y la calificación final.

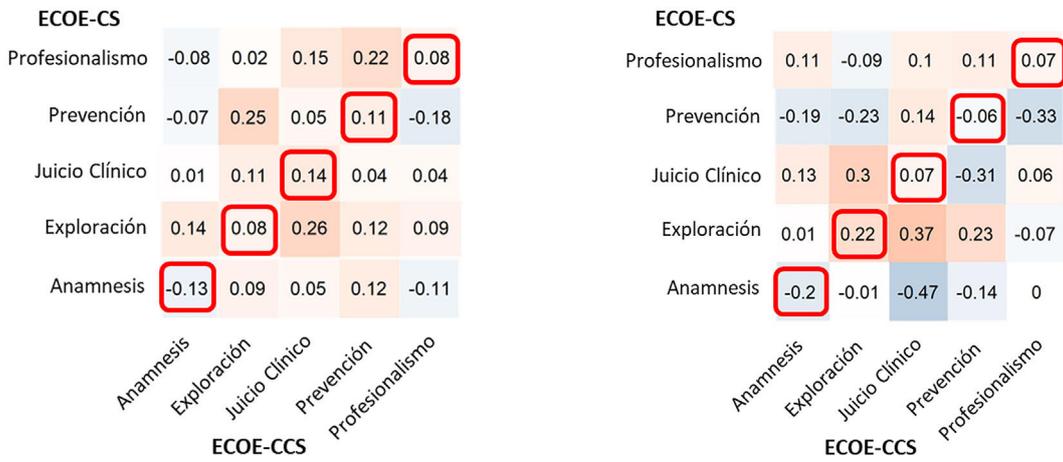
Se analizó la consistencia interna de cada prueba respecto a sus dominios competenciales y el alfa de Cronbach fue: ECOE-CS-1  $\alpha=0,62$ ; ECOE-CS-2  $\alpha=0,54$  y ECOE-CCS  $\alpha=0,38$ .

Se valoró la posible influencia en las calificaciones de las puntuaciones asignadas a los ítems. Las ECOE-CS-1 y 2 fueron homogéneas en el número medio de puntos por ítem pero, comparativamente, la puntuación media de los ítems de la ECOE-CCS fue superior en anamnesis (2,86 puntos/ítem vs. 2,22 en ECOE-CS), exploración clínica (3,74 puntos/ítem vs. 2,63 en ECOE-CS) y prevención y promoción de la salud (3,31 puntos/ítem vs. 3,22 en ECOE-CS). Por el contrario, la puntuación media de ítems fue superior en la ECOE-CS en juicio clínico (3,62 puntos/ítem vs. 3,39 en ECOE-CCS) y profesionalismo (3,81 puntos/ítem vs. 3,44 en ECOE-CCS).

Se analizó la correlación de resultados de todos los alumnos en las 2 pruebas, por calificación final y cada dominio competencial. Además, se analizó en el subgrupo de alumnos que mejores calificaciones había alcanzado en cada uno de los dominios competenciales de la ECOE-CS (Q4,

**Tabla 3** Resultados globales y por dominios competenciales de los alumnos, y significación de la comparación (% de puntos alcanzados; media  $\pm$  DT)

	Media ECOE-CS (N=108)	Media ECOE-CCS (N=108)	Comparativa ECOE-CS vs. ECOE-CCS	Media ECOE-CS-1 (N=72)	Media ECOE-CS-2 (N=36)	Comparativa ECOE-CS-1 vs. ECOE-CS-2
Anamnesis	71,2 $\pm$ 7,9	72,3 $\pm$ 4,1	0,191	69,9 $\pm$ 7,7	73,7 $\pm$ 7,9	0,023
Exploración clínica	75,6 $\pm$ 7,5	85,8 $\pm$ 7,9	<0,001	76,2 $\pm$ 7,3	74,4 $\pm$ 8,0	0,270
Habilidades técnicas	80,2 $\pm$ 11,4			81,0 $\pm$ 11,1	78,7 $\pm$ 11,8	0,341
Comunicación	76,8 $\pm$ 8,6			77,5 $\pm$ 8,8	75,3 $\pm$ 8,3	0,196
Juicio clínico	64,2 $\pm$ 10,0	68,2 $\pm$ 5,9	<0,001	62,1 $\pm$ 10,4	68,5 $\pm$ 7,6	<0,001
Prevención y promoción	52,6 $\pm$ 11,1	64,0 $\pm$ 10,1	<0,001	52,5 $\pm$ 11,3	52,7 $\pm$ 10,9	0,920
Relaciones interprofesionales	60,5 $\pm$ 8,9			58,7 $\pm$ 8,9	64,3 $\pm$ 7,8	0,001
Profesionalismo	62,9 $\pm$ 13,3	76,3 $\pm$ 19,4	<0,001	58,7 $\pm$ 12,1	71,3 $\pm$ 11,4	<0,001
Puntuación total	69,5 $\pm$ 5,0	71,4 $\pm$ 4,3	0,005	68,6 $\pm$ 5,1	71,3 $\pm$ 4,4	0,006



**Figura 1** Coeficiente de correlación de Pearson (R) entre dominios competenciales de la ECOE-CS y la ECOE-CCS para todos los alumnos (izquierda) y alumnos Q4 (derecha). Se enmarcan los valores de R entre dominios competenciales homónimos.

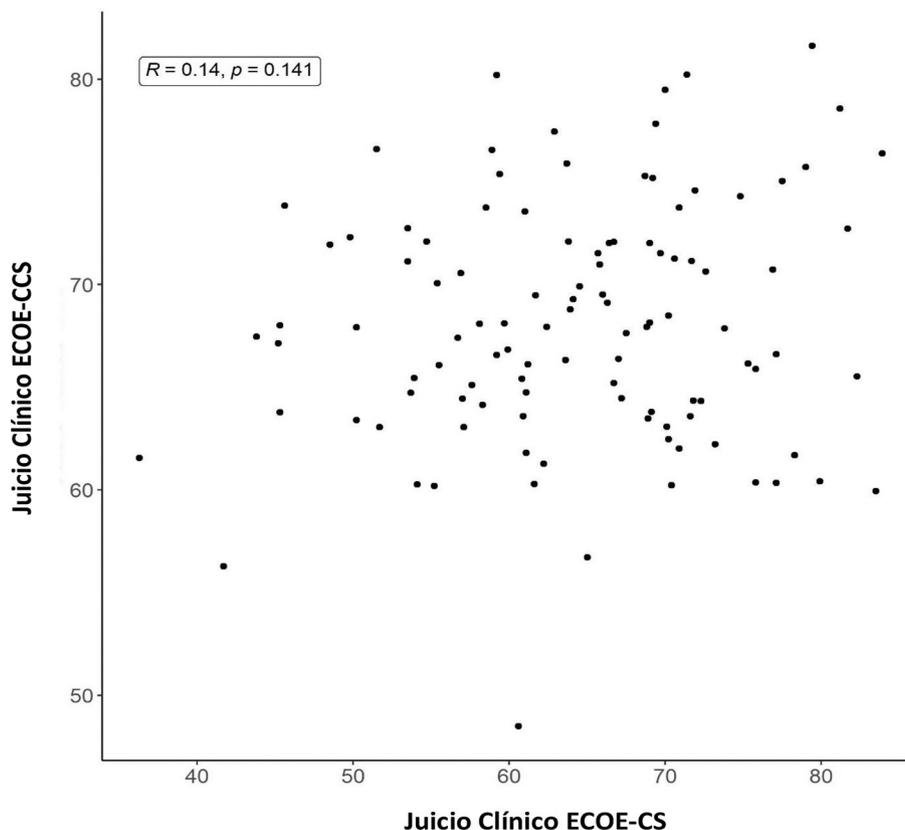
percentil  $\geq 75$ ), para buscar un hipotético mejor rendimiento en la ECOE-CCS de estos alumnos en esos mismos dominios. No se demostró correlación entre calificaciones globales de la ECOE-CS con la ECOE-CCS:  $R=0,184$  ( $p=0,056$ ). Tampoco se obtuvo correlación al analizarlas por separado: ECOE-CS-1 vs. ECOE-CCS:  $R=0,205$  ( $p=0,084$ ) y ECOE-CS-2 vs. ECOE-CCS:  $R=0,140$  ( $p=0,414$ ).

En la **figura 1** se muestran los valores del coeficiente de correlación de Pearson entre los mismos dominios competenciales de las 2 pruebas, para todos los alumnos y

para el subgrupo de alumnos Q4, sin que se demuestre correlación significativa en ninguno de ellos ( $p$  entre 0,141 y 0,764).

En la **figura 2** se presenta, como ejemplo, la dispersión de resultados obtenidos por todos los alumnos en juicio clínico en ambas pruebas, sin correlación.

Se estudió también la existencia de concordancia de cada alumno en los mismos dominios competenciales dentro de las distintas estaciones de la ECOE-CS e, igualmente, los distintos casos clínicos de la ECOE-CCS, sin demostrarse



**Figura 2** Coeficiente de correlación de Pearson (R) entre juicio clínico de la ECOE-CS y la ECOE-CCS (todos los alumnos).

concordancia en ningún caso ( $R = 0 - 0,62$  en ECOE-CS;  $R = 0 - 0,23$  en ECOE-CCS).

## Discusión

La evaluación de la competencia es un proceso complejo, que requiere múltiples pruebas y evaluadores, y metodologías válidas y fiables<sup>8,9</sup>. Evaluar la competencia (*competence*), no obstante, difiere de evaluar el desempeño o actuación habitual en condiciones reales (*performance*)<sup>4,10</sup>. Entre los métodos de evaluación de la competencia, la ECOE es una herramienta capaz de evaluar el desempeño en entorno de simulación<sup>11</sup>.

En la literatura existen diversas herramientas tecnológicas para evaluar la competencia, mejorando la accesibilidad y tratamiento de los datos, y reduciendo costos de producción<sup>12</sup>. Entre ellas, existen programas informáticos de simulación de casos clínicos por ordenador (Computer-based Case Simulations, CCS), también denominados pacientes virtuales (PV), creados para obviar las dificultades logísticas de la ECOE<sup>13</sup>. Sus inconvenientes son los altos costes y las limitaciones de interacción física y de fidelidad a la realidad<sup>12,14</sup>. El Step-3 CCS es un ejemplo de ello: despliega 13 casos clínicos mediante el *software* de simulación virtual *Primum CCS*<sup>15</sup>. El Step-3 CCS incluye, además de la evaluación con dicho *software*, más de 400 preguntas de opción múltiple administradas en 2 días distintos<sup>16</sup>.

La ECOE *online* (*Virtual ECOE*, *VECOE*) es otra herramienta tecnológica cuya principal característica es que conserva la interacción con pacientes simulados en tiempo real a través de plataformas *online*<sup>17,18</sup>.

También existen experiencias de evaluación con preguntas de opción múltiple sobre casos clínicos por ordenador (*Computer-based Cases Test*, *CCT*)<sup>19</sup>.

La ECOE-CCS se asimiló con los casos clínicos virtuales del Step-3 pero, a diferencia de éste, no utiliza un *software* simulador de casos clínicos CCS, sino que se trata de un examen de casos clínicos por ordenador (CCT). En coherencia con la literatura internacional y para evitar confusiones sobre su metodología evaluativa, no se debería denominar ECOE-CCS, sino prueba CCT, al asimilarse a dicha metodología, y no a la ECOE convencional.

Una limitación de nuestro estudio es que la muestra corresponde a alumnos de una única facultad y los resultados podrían atribuirse a características propias, distintas a la media de otras facultades. Sin embargo, los resultados globales de nuestros alumnos en la ECOE-CS son similares a los descritos por otras facultades nacionales<sup>20-22</sup> y los resultados de la ECOE-CCS, desprendidos de datos internos analizados por el Comité ECOE-CND, también nos sitúan en un rango similar respecto al resto de facultades españolas.

Otra limitación es que la ECOE-CS, por el contexto COVID-19, solo tuvo 12 estaciones y esto pudo afectar a los resultados de consistencia interna. Por el contrario, creemos que favoreció mejores condiciones de comparación con la ECOE-CCS, por similar número de casos e ítems.

La diferencia de resultados entre las ECOE-CS-1 y 2 pudiera deberse a los distintos grados de dificultad de algunas estaciones específicas y al menor número de alumnos en la ECOE-CS-2 y no tanto a su capacitación, ya que la asignación fue aleatoria. Los mejores resultados en los turnos de tarde parecen confirmar que pudo existir un

efecto copia en ambas versiones. Esta hipótesis se corrobora porque, de los 52 alumnos (48%) que cumplimentaron una encuesta anónima de satisfacción al finalizar la prueba, 22 (42%) reconocieron haber recibido información de compañeros durante el día.

La ponderación de los ítems también pudo influir en los resultados finales de las 2 pruebas, si bien consideramos que no justifica la falta de correlación entre dominios competenciales pues las diferencias de puntuación en los ítems no suponen cambios directamente proporcionales en los resultados de los alumnos, que pueden acertarlos o no, y alumnos brillantes en sus resultados ECOE-CS obtuvieron malos resultados en la ECOE-CCS y a la inversa.

La cohesión entre dominios competenciales ( $\alpha$  de Cronbach) se demostró mayor en las ECOE-CS que en la ECOE-CCS, aún sin ofrecer una consistencia interna robusta ( $\alpha = 0,38$  a  $0,62$ ). No obstante, los valores superiores son similares a los descritos en una revisión sistemática del tema y atribuyen una mayor fiabilidad a un número mayor de estaciones, la presencia de segundos evaluadores y algunos factores no conocidos<sup>23</sup>.

En nuestro estudio no se ha podido demostrar correlación de resultados globales ni por dominios competenciales entre ECOE-CS y ECOE-CCS. Existen 2 estudios donde se obtiene correlación, pero se trata de un *software* de pacientes virtuales tipo CCS. En el primero, con un programa web que procesaba lenguaje natural y ofrecía respuestas coherentes en las distintas fases del caso, se demostró correlación global, pero no por competencias<sup>24</sup>. En el segundo, con algunos déficit metodológicos no aclarados, se correlacionan resultados en algunas competencias del *software* con la calificación global de la ECOE<sup>25</sup>.

En otro estudio se comparó un examen de casos clínicos por ordenador tipo CCT con la competencia frente a pacientes simulados ocultos durante la consulta médica y se demostró una correlación inversa de resultados entre ambas pruebas<sup>19</sup>. Los autores concluyeron que la evaluación con pacientes simulados y casos clínicos por ordenador no rinden similar información sobre la competencia y/o desempeño del candidato, al evaluar distintas dimensiones y tener constructos diferentes.

En un estudio más similar al nuestro, 80 alumnos realizaron un examen de 8 casos clínicos computarizados donde se medían conocimientos y habilidades de toma de decisiones con preguntas verdadero-falso y de opción múltiple, y una ECOE de 4 estaciones y 4 informes clínicos posteriores, evaluada por observadores externos a los pacientes estandarizados. Se demostró correlación débil ( $R = 0,32$  a  $0,47$ ) tan solo para los resultados finales de los alumnos, no por competencias<sup>26</sup>.

La ausencia de correlación de resultados por alumno en las mismas competencias entre estaciones independientes de las ECOE-CS, creemos que es coherente con el conocimiento actual de que es necesario un determinado número de estaciones ECOE para garantizar la validez y fiabilidad de los resultados y no es posible inferirlos sólo con un número inferior de estaciones, pues los dominios competenciales se evalúan por los resultados del conjunto de las estaciones<sup>4,5,23</sup>. No podemos interpretar el significado de la falta de correlación interna entre casos clínicos de la ECOE-CCS por no existir literatura ni experiencia previa sobre las dimensiones competenciales a medir.

La validez de una prueba, o capacidad de medir adecuadamente aquello para lo que se diseñó, es el fundamento de la evaluación y requiere un largo y elaborado proceso de construcción que pasa, entre otros, por buscar la correlación de resultados con otras pruebas que midan lo mismo<sup>8,9,11,27-29</sup>. En nuestro estudio esta correlación no ha existido. Mientras que en la ECOE ya se ha medido su validez y fiabilidad, por el momento no se conocen las de la mal denominada ECOE-CCS.

El desempeño frente a un paciente simulado (EEOE-CS) y el razonamiento ejercitado en un examen por ordenador sobre unos dominios competenciales en relación a un caso clínico (prueba CCT, erróneamente denominada ECOE-CCS) suponen procesos psicológicos distintos que deben medirse con herramientas evaluativas de constructo psicométrico distinto. Esto es especialmente evidente cuando, por ejemplo, se evalúa la competencia en exploración clínica frente a un paciente o a un caso clínico por ordenador. Ambas pruebas evalúan competencias distintas y en diferentes niveles de la pirámide de Miller por lo que no es posible asimilarlas ni ponderarlas como iguales<sup>8,10,30</sup>.

Tal y como definía Harden, «la ECOE es un examen clínico o práctico. Es una evaluación del desempeño y se centra más en lo que los alumnos pueden hacer que en lo que saben. Lo que se evalúa es lo que hace el examinado cuando se enfrenta a un paciente o a una situación, no lo que sabe o las respuestas que escribe a una pregunta teórica sobre el tema»<sup>2</sup>.

La mal denominada ECOE-CCS es un formato novedoso de examen en nuestras facultades que pudiera tener un papel en la evaluación del alumno, pero no debe correlacionarse ni confundirse con la ECOE, ni sus resultados por competencias deben ponderarse con las competencias de una ECOE. Serán necesarios más estudios comparativos entre pruebas que incluyan resultados del conjunto de facultades para ratificar o no las hipótesis de este trabajo y extraer conclusiones para la definición y mejora de esta prueba en el futuro.

## Responsabilidades éticas

El protocolo del estudio fue aprobado por el Comité de Ética de la Investigación de la Universidad Francisco de Vitoria con fecha 21 de diciembre de 2022.

En relación a la exención de consentimiento informado de los participantes, deseamos exponer que la aclaración sobre este punto nos fue exigida previamente en el protocolo presentado al Comité de Ética para su aprobación.

Los datos recogidos para el estudio no contienen datos personales, tan solo el resultado de las evaluaciones de alumnos sin ningún dato que permita la asociación del participante con su identidad. Además, todos los datos recogidos en el transcurso del estudio fueron tratados de forma estrictamente confidencial. Los datos solo han sido utilizados para desarrollar el estudio y su tratamiento se ha destinado estrictamente a fines estadísticos. El equipo investigador se comprometió individualmente, con documento firmado manuscrito, a mantener la identidad y datos de los participantes en el anonimato, y a omitir cualquier acción encaminada a la identificación de los participantes del estudio. Por todo ello, el Comité de Ética aceptó la exención del consentimiento informado.

## Conflicto de intereses

Los autores declaran que no tienen intereses económicos ni relaciones personales que pudieran haber influido en el trabajo presentado en este artículo.

## Agradecimientos

A Roger Ruiz Moral y Santiago Álvarez Montero por sus acertadas orientaciones en la redacción del manuscrito.

## Bibliografía

- Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *Br Med J* [Internet]. 1975;1(5955):447-51 [consultado 23 Oct 2021]. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1672423/>.
- Harden RM. What is an OSCE? *Med Teach*. 1988;10(1):19-22.
- Khan KZ, Gaunt K, Ramachandran S, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part II: organisation & administration. *Med Teach*. 2013;35(9) e1447-1463.
- Epstein RM. Assessment in medical education. *N Engl J Med*. 2007;356(4):387-96.
- Sureda-Demeulemeester E, Ramis-Palmer C, Sesé-Abad A. La evaluación de competencias en Medicina. *Rev Clínica Esp* [Internet]. 2017;217(9):534-42 [consultado 23 Jun 2022]. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0014256517301479>.
- García-Estañ López J. Prueba Nacional de Evaluación de Competencias Clínicas de la Conferencia Nacional de Decanos de Facultades de Medicina de España. *FEM*. 2013;16(Supl 3) S1-70.
- García-Seoane JJ, Ramos-Rincón JM, Lara-Muñoz JP, López-Román A, Cardoner-Álvarez N, Lorenzo-Peñuelas A, et al. Cambios en el examen clínico objetivo y estructurado (EEOE) de las facultades de Medicina durante la COVID-19. Experiencia de una ECOE de casos-clínicos computarizados simulados (EEOE-CCS) conjunta. *Rev Clínica Esp* [Internet]. 2021 [consultado 20 May 2021]; Disponible en: <https://www.sciencedirect.com/science/article/pii/S0014256521000527>.
- 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* [Internet]. 2017;39(6):609-16 [consultado 23 Oct 2021]. Disponible en: <https://www.tandfonline.com/doi/full/10.1080/0142159X.2017.1315082>.
- Boulet JR, Durning SJ. What we measure ... and what we should measure in medical education. *Med Educ* [Internet]. 2019;53(1): 86-94 [consultado 23 Oct 2021]. Disponible en: <https://onlinelibrary.wiley.com/doi/10.1111/medu.13652>.
- 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 Médica* [Internet]. 2016;17(4):130-9 [consultado 23 Jun 2022]. Disponible en: <https://www.sciencedirect.com/science/article/pii/S157518131630078X>.
- Khan KZ, Ramachandran S, Gaunt K, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part I: an historical and theoretical perspective. *Med Teach*. 2013;35(9) e1437-1446.
- Association of American Medical Colleges. Effective Use of Educational Technology in Medical Education: Summary Report of the 2006 AAMC Colloquium on Educational Technology. Washington, DC:AAMC; 2007;7 [Internet]. 2007 [consultado 24 Oct 2021]. Disponible en: <https://presentationdesignformeded>.

- [files.wordpress.com/2013/05/effective-use-of-educational-technology-in-medical-education.pdf](https://files.wordpress.com/2013/05/effective-use-of-educational-technology-in-medical-education.pdf).
13. Ward RC, Muckle TJ, Kremer MJ, Krogh MA. Computer-based case simulations for assessment in health care: a literature review of validity evidence. *Eval Health Prof* [Internet]. 2019;42(1):82–102 [consultado 23 Oct 2021]. Disponible en: <http://journals.sagepub.com/doi/10.1177/0163278717718609>.
  14. Fink MC, Reitmeier V, Stadler M, Siebeck M, Fischer F, Fischer MR. Assessment of diagnostic competences with standardized patients versus virtual patients: Experimental study in the context of history taking. *J Med Internet Res* [Internet]. 2021;23(3) [consultado 12 Nov 2022]. e21196. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7974754/>.
  15. United States Medical Licensing Examination. Computer-based Case Simulations | USMLE [Internet] [consultado 24 Oct 2021]. Disponible en: <https://www.usmle.org/step-3-test-question-formats/computer-based-case-simulations2021>.
  16. United States Medical Licensing Examination. Step 3 Exam Content | USMLE [Internet] [consultado 13 Oct 2021]. Disponible en: <https://www.usmle.org/step-exams/step-3/step-3-exam-content2021>.
  17. Blythe J, Patel NSA, Spiring W, Easton G, Evans D, Meskevicus-Sadler E, et al. Undertaking a high stakes virtual OSCE (“VOSCE”) during Covid-19. *BMC Med Educ* [Internet]. 2021;21(1):221 [consultado 18 Nov 2021]. Disponible en: <https://doi.org/10.1186/s12909-021-02660-5>.
  18. Hopwood J, Myers G, Sturrock A. Twelve tips for conducting a virtual OSCE. *Med Teach* [Internet]. 2021;43(6):633–6 [consultado 17 Nov 2021]. Disponible en: <https://doi.org/10.1080/0142159X.2020.1830961>.
  19. Schuwirth L, Gorter S, Van der Heijde D, Rethans JJ, Brauer J, Houben H, et al. The role of a computerised case-based testing procedure in practice performance assessment. *Adv Health Sci Educ Theory Pract*. 2005;10(2):145–55.
  20. Ramos JM, Martínez-Mayoral MA, Sánchez-Ferrer F, Morales J, Sempere T, Belinchón I, et al. Análisis de la prueba de evaluación clínica objetiva estructurada (ECO) de sexto curso en la Facultad de Medicina de la Universidad Miguel Hernández de Elche. *Educ Médica* [Internet]. 2019;20:29–36 [consultado Oct 2021]. Disponible en: <https://www.sciencedirect.com/science/article/pii/S1575181317301584>.
  21. García-Puig J, Vara-Pinedo F, Vargas-Núñez JA. Implantación del examen clínico objetivo y estructurado en la Facultad de Medicina de la Universidad Autónoma de Madrid. *Educ Médica* [Internet]. 2018;19(3):178–87 [consultado 10 Oct 2021]. Disponible en: <https://www.sciencedirect.com/science/article/pii/S1575181317300232>.
  22. Kronfly Rubiano E, Ricarte Díez JI, Juncosa Font S, Martínez Carretero JM. Evaluación de la competencia clínica de las facultades de medicina de Cataluña, 1994-2006: evolución de los formatos de examen hasta la evaluación clínica objetiva y estructurada (ECO). *Med Clínica* [Internet]. 2007;129(20):777–84 [consultado 10 Oct 2021]. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0025775307729937>.
  23. Brannick MT, Erol-Korkmaz HT, Prewett M. A systematic review of the reliability of objective structured clinical examination scores. *Med Educ* [Internet]. 2011;45(12):1181–9 [consultado 6 Mar 2023]. Disponible en: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2923.2011.04075.x>.
  24. Oliven A, Nave R, Gilad D, Barch A. Implementation of a web-based interactive virtual patient case simulation as a training and assessment tool for medical students. *Stud Health Technol Inform*. 2011;169:233–7.
  25. Fida M, Kassab SE. Do medical students’ scores using different assessment instruments predict their scores in clinical reasoning using a computer-based simulation? *Adv Med Educ Pract* [Internet]. 2015;6:135–41 [consultado 17 Nov 2021]. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4345894/>.
  26. Guagnano MT, Merlitti D, Manigrasso MR, Pace-Palitti V, Sensi S. New medical licensing examination using computer-based case simulations and standardized patients. *Acad Med J Assoc Am Med Coll*. 2002;77(1):87–90.
  27. American Educational Research Association. Standards for educational and psychological testing. [Internet]#;59].1999.a ed#;59]. #;59]; 2005;9–24. Washington DC; [consultado 6 Mar 2023]. Disponible en: [https://www.aera.net/Portals/38/1999%20Standards\\_revised.pdf](https://www.aera.net/Portals/38/1999%20Standards_revised.pdf). Washington DC; [consultado 6 Mar 2023]. Disponible en.
  28. Tavakol M, Dennick R. The foundations of measurement and assessment in medical education. *Med Teach* [Internet]. 2017;39(10):1010–5 [consultado 6 Mar 2023]. Disponible en: <https://doi.org/10.1080/0142159X.2017.1359521>.
  29. Downing SM. Validity: on the meaningful interpretation of assessment data. *Med Educ* [Internet]. 2003;37(9):830–7 [consultado 6 de Mar 2023]. Disponible en: <https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2923.2003.01594.x>.
  30. Schuwirth LWT, CPM Van Der Vleuten. The use of clinical simulations in assessment. *Med Educ* [Internet]. 2003;37(s1):65–71 [consultado 17 Nov 2021]. Disponible en: <https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2923.37.s1.8.x>.