

Confiabilidad y validez de las listas de cotejos del Examen Clínico Objetivo Estructurado para el aprendizaje por competencias de Cirugía

Mario A. Chávez Gil¹, Moisés Barrantes Cabrera²

¹ Departamento de Cirugía, Universidad Nacional de Trujillo; mariochg@gmail.com

² Escuela de Postgrado, Universidad Nacional de Trujillo; moisesbarrantesc@hotmail.com

Recibido: 21-05-2014

Aceptado: 26-09-2014

RESUMEN

En el presente trabajo se ha evaluado el aprendizaje por competencias de cirugía de estudiantes del sexto año de medicina de la Universidad Nacional de Trujillo, Perú, así como la validez y confiabilidad de las listas de cotejos del examen clínico objetivo estructurado (ECO). La investigación se llevó a cabo utilizando listas de cotejos para diez estaciones del ECOE sobre cirugía general, neurocirugía, oftalmología, otorrinolaringología y urología. Dos estaciones de cirugía general utilizaron paciente simulador. En neurocirugía una estación utilizó una placa de tomografía axial computarizada y la segunda requirió evaluación con escala de Glasgow. En urología una estación utilizó un paciente simulador para cólico renal y la segunda una maqueta para colocación de sonda de Foley. En oftalmología, una estación utilizó una imagen de ojo rojo y la segunda un paciente simulador; en otorrinolaringología, una estación utilizó una tomografía axial computarizada y la segunda un paciente simulador para valoración de prueba auditiva con diapason. Se encontró que: el rendimiento por competencias quirúrgicas de los 86 estudiantes de medicina del sexto año de la Universidad Nacional de Trujillo fue muy satisfactorio (calificativo promedio: 16 puntos); de las 10 listas de cotejos de las pruebas de ECOE, en seis estaciones la confiabilidad alcanzó niveles de 0,33 a 0,90 y cuatro no tuvieron confiabilidad; en cuanto a la validez, cuatro estaciones tuvieron listas de cotejos válidas y seis estaciones tuvieron listas de cotejos imprecisas.

Palabras clave: Validez de una prueba, confiabilidad de una prueba, listas de cotejo, ECOE.

ABSTRACT

In this research work we have evaluated the learning by competences of sixth year medicine students of "Universidad Nacional de Trujillo", Perú, as well as the validity and reliability of the matching checklists of the Objective Structured Clinical Examination (OSCE). The research was carried out by using ten matching checklists comparisons OSCE stations on general surgery, neurosurgery, ophthalmology, otolaryngology and urology. Two stations of general surgery used patient simulator. In Neurosurgery a station used a computerized tomography plate of a case of traumatic brain injury and the second one required an assessment with a Glasgow scale. In urology a station used a patient simulator for renal colic and the second one used a support for Foley catheter placement. In ophthalmology a station used a red-eye image and the second one used a patient simulator; in otorhinolaryngology, a station used an axial computerized tomography scan and the second one a patient simulator for evaluation of hearing with Rinne test. It was found that: the performance of surgical competences of the 86 medical students in the sixth year of the "Universidad Nacional de Trujillo" of Trujillo was very satisfactory (the average qualification was 16 points); from the ten matching checklists of the 10 comparisons of OSCE tests, in six OSCE stations the reliability reached levels from 0,33 to 0,90 and four matching checklists had no reliability; as regards to the validity of the tests, four stations had valid matching checklists and six ones had inaccurate matching checklists.

Keywords: validity of a test, reliability of a test, checklists, OSCE.

I. INTRODUCCIÓN

Actualmente, con la globalización las universidades tratan de uniformizar criterios de enseñanza aprendizaje y de evaluación como por ejemplo, el Proyecto Tuning Europa (González, 2003), Declaración de Bolonia (1999) y el Proyecto Tuning América Latina (Beneitone et al, 2007: 256).

Las actuales demandas de la sociedad en la formación de sus profesionales intentan ser satisfechas por un nuevo paradigma: el de la Educación Basada en Competencias (Andrade, 2008: 8; Jones et al, 2001: 699-703; Miller, 1990: 563-567; Reta et al, 2006: 75-83, Ross et al, 1988: 45-46 y Wass et al, 2001: 945-49).

Entre los instrumentos de evaluación de competencias surge el Examen Clínico Objetivo Estructurado, descrito por primera vez en Escocia en el año 1975, por Harden (1975: 447-451). Desde entonces, este sistema estructurado de evaluación ha tenido mucha aceptación (Cuschieri, 1979: 400-5, Jones et al, 2001: 699-703; Joorabchi, 1991: 757-62; Miller, 1990: 563-567; Petrusa, 1990: 573-7; Reta et al, 2006: 75-83; Ross et al., 1988: 45-46; Wass et al, 2001: 945-949; Sloan, 1993: 343-50).

La evaluación del aprendizaje se realiza de una manera formativa (Sloan et al, 1995: 735-742; Escuela Andaluza de Salud Pública, 2008).

Durante las prácticas hospitalarias, se evalúa el progreso y se identifican las áreas y competencias susceptibles de mejora, facilitando la reflexión y la retroalimentación, se evalúan los aspectos procedimental, cognitivo y actitudinal.

La evaluación de las competencias profesionales es cada vez mayor en la educación superior y por ello, en la educación médica.

Las competencias son el conjunto de conocimientos, habilidades, actitudes y valores que interrelacionados conducen hacia un desempeño profesional eficiente (Universidad Nacional Autónoma de México, s/f).

La evaluación incluye varios procedimientos para conseguir información sobre el aprendizaje de los estudiantes, lo que permite a los profesores emitir juicios de valor y adoptar las estrategias para el mejoramiento del proceso enseñanza-aprendizaje en la formación de los estudiantes y en beneficio de la sociedad (Sánchez, 2012: 19).

Los complejos procedimientos seguidos por los diseñadores profesionales de pruebas, son poco conocidos por la mayoría de los profesores; con frecuencia, éstos consideran que los exámenes son algo desagradable adjunto a la enseñanza en lugar de verlos como parte integral y formativa del proceso educativo (Aiken, 2003: 18-19).

Las *listas de cotejos* utilizadas en la valoración de conductas y habilidades de los alumnos, requieren ser válidas y confiables, por este motivo es que fueron sometidas a pruebas estadísticas para determinar si contaban con tales atributos.

La confiabilidad es la capacidad de un examen de arrojar un resultado consistente cuando se repite en ocasión diferente por las mismas personas. Se expresa como un coeficiente de correlación, siendo 1.0 la correlación perfecta y cero ninguna correlación. Mientras más alta es la cifra de confiabilidad, es mayor su peso como validez.

La validez de una prueba nos permite determinar si mide lo que se propone medir y la clasificación según la correlación pregunta-prueba, la misma que se calcula con el puntaje total de la prueba. Según nos indica Bazán (2000: 5), es:

CLASIFICACIÓN	ÍNDICE DE VALIDEZ
Muy buena	0,20 – 1
Aceptable	0 – 0,19
Eliminar	< 0

En la Figura 1, se ilustran los conceptos de validez y confiabilidad de un instrumento de evaluación, usando la analogía de un tiro al blanco según Bostwick y Kyte (2005), citados por Hernández (2006: 284-85).



Fig. 1. Validez vs confiabilidad.

En el Departamento de Cirugía de la Facultad de Medicina de la Universidad Nacional de Trujillo, (Universidad Nacional de Trujillo, 2011) el Sílabo de Cirugía II se encuadra en el paradigma de la enseñanza-aprendizaje por competencias. Al finalizar el curso, se evalúan los resultados de las competencias obtenidos por los alumnos al término de su rotación.

La experiencia curricular teórico - práctica para los estudiantes de sexto año de Medicina es parte de la formación médico-quirúrgica en lo referente al diagnóstico y tratamiento inicial actualizado de las patologías médico-quirúrgicas de las especialidades de Cirugía General, Neurocirugía, Oftalmología, Otorrinolaringología y Urología.

Para el estudiante el curso dura cuarenta días siendo la extensión de 24 horas de horas teóricas y 96 horas de práctica.

Las competencias adquiridas por los estudiantes son: generales y específicas. El estudiante debe saber qué hacer y cómo hacer en el lugar y momento adecuado (Lozano, 2010) y con un compromiso e identificación con su trabajo.

Las competencias generales son:

- Elaboración de la historia clínica, establecer hipótesis diagnósticas, solicitar exámenes auxiliares necesarios, prescripción de medicamentos y realizar procedimientos terapéuticos bajo la supervisión de los profesores.
- Reconocer los problemas de salud más frecuentes de la clínica quirúrgica en las áreas mencionadas líneas arriba.
- Intervenir en el examen y evaluación del paciente en las áreas mencionadas.
- Realizar en el laboratorio de técnica quirúrgica, los procedimientos quirúrgicos más comunes.

Las competencias específicas son las siguientes:

Desarrollo de conocimientos, habilidades, destrezas y comportamiento adecuado en el manejo inicial de las principales patologías en: Oftalmología (Ametropía-ambliopía, estrabismo, conjuntivitis, catarata, glaucoma y retinopatía diabética), Neurocirugía (Traumatismo encéfalo-craneano, traumatismo vértebro-medular, hemorragia sub-aracnoidea, hipertensión endo-craneana y lesiones mielo-radiculares), Urología (Hiperplasia benigna de la próstata, cáncer de próstata, incontinencia urinaria y litiasis urinaria), Cirugía General (Abdomen agudo, hernias, eventraciones, apendicitis aguda, poliposis colónica, enfermedad hemorroidal, abscesos hepáticos, quiste hepático, hidatidosis, pancreatitis aguda, colecistitis y colangitis aguda), Otorrinolaringología (Sinusitis, rinitis alérgica, resfrío común, vestibulitis nasal, fractura nasal, mandibular, malar y maxilar superior, otitis, estomatitis, amigdalitis, epistaxis).

Al final de cada rotación, el estudiante es sometido a evaluación sumativa de las diferentes actividades lectivas: clases, técnica quirúrgica, casos clínicos, seminarios y monografías. Se realiza un examen escrito que evalúa conocimientos y se evalúan habilidades y competencias mediante el examen clínico objetivo estructurado (ECO). La evaluación sumativa permite determinar la aptitud

para ejercer (Sociedad Española de Médicos de Atención Primaria, 2009). La evaluación se rige por la calificación vigesimal. La nota aprobatoria mínima es 10.5.

La valoración del aprendizaje de la experiencia curricular se rige por la siguiente ponderación:

Prácticas hospitalarias	20%
Técnicas quirúrgicas	5%
Caso clínico	10%
Clase teórica	5%
Seminario	5%
Monografía	5%
ECOЕ	10%
Examen escrito	40%
Peso total	100%

En el Departamento de Cirugía, se está cambiando el paradigma clásico de enseñanza aprendizaje centrada en la enseñanza por el nuevo paradigma de la educación centrada en el aprendizaje, en la cual el objetivo es el estudiante como generador de sus propios conocimientos y a la vez gestor de los mismos. El profesor necesariamente tiene que ir cambiando su rol, siendo más orientador de la organización del aprendizaje.

Dentro de este contexto, en la Facultad de Medicina se viene efectuando la evaluación clínica objetiva estructurada (ECOЕ) como la valoración de habilidades y competencias de los estudiantes del Departamento de Cirugía. Esta evaluación se desarrolla desde 1996, hubo una ligera experiencia previa el año 1995 en el Departamento de Medicina pero el examen fue desactivado el año siguiente. Es así que los alumnos de 6° Año del 2011 que cursaban las especialidades de Cirugía General, Neurocirugía, Otorrinolaringología, Oftalmología y Urología fueron evaluados con el formato ECOЕ, que incorpora diversos instrumentos evaluativos y se desarrolla a lo largo de sucesivas estaciones que simulan situaciones clínicas.

La lista de cotejos es una relación de características, aspectos, cualidades, etc. sobre las que interesa determinar su presencia o ausencia. Se centra en registrar la aparición o no de una conducta durante la evaluación. Es un instrumento que forma parte de la técnica de observación, es decir, a través de su aplicación se obtiene información sobre conductas y acontecimientos habituales de los estudiantes y permite conocer aspectos como intereses, actitudes, habilidades, destrezas, etc.

Las pruebas de evaluación deben tener validez y confiabilidad (Miller 1990: 563-567) (Sloan et al. 1996: 225-30).

Para el examen de fin de curso en las escuelas de medicina es deseable que la confiabilidad sea de 0,70 a 0,89. Para las evaluaciones más altas como los exámenes de certificación o licencia en medicina, se espera que sea de al menos 0.90. Esto es referencial ya que existen diferencias de opinión entre los expertos (Downing 2004: 1006-1012).

El ECOЕ es un formato de examen, una excelente herramienta de evaluación de destrezas o habilidades.

El propósito del presente trabajo fue evaluar el aprendizaje por competencias de cirugía de los estudiantes del sexto año de medicina, así como valorar el nivel de validez y confiabilidad de las listas de cotejos de las estaciones del ECOЕ, las mismas que fueron preparadas por los profesores del departamento de Cirugía de la Facultad de Medicina, Universidad Nacional de Trujillo, Perú.

II. MATERIAL Y METODOS

2.1 Objeto de estudio

La presente investigación tuvo como objeto de estudio a estudiantes de 6° Año de Medicina de la Universidad Nacional de Trujillo (UNT), Perú, en relación a las llamadas *listas de cotejos* que se utilizan para determinar el grado de competencias alcanzado por los estudiantes durante las estaciones del Examen Clínico Objetivo Estructurado, ECOЕ.

Las listas de cotejos, elaboradas por los profesores del departamento de Cirugía de la Facultad de medicina de la UNT, permiten observar y registrar las conductas, actitudes, habilidades y destrezas

de los alumnos. Las listas de cotejos contienen ítems de calificación de acuerdo a si era correcta o no una observación (Ver Anexo para un ejemplo).

Durante el estudio 18 docentes realizaban labores con los estudiantes de 6° Año en las áreas de Cirugía General, Neurocirugía, Oftalmología, Otorrinolaringología y Urología. En cada especialidad, los docentes prepararon dos estaciones.

En Cirugía General, las estaciones fueron: anamnesis de dolor abdominal y examen físico de abdomen, ambas contaron con paciente simulador.

En Urología, las estaciones fueron: colocar sonda de Foley y examen físico en cólico renal, la primera utilizaba una maqueta y la segunda utilizaba un paciente simulador.

En Neurocirugía, una estación consistió en observar una placa de tomografía axial computarizada, en la que el estudiante describía los hallazgos y luego anotaba el diagnóstico correcto; la segunda estación consistía en valorar la escala de Glasgow en un paciente en coma, para determinar el estado de conciencia.

En Otorrinolaringología una estación consistía en describir hallazgos y determinar el diagnóstico en una placa radiográfica de senos para nasales y la otra consistía en realizar una prueba de audición de Rinne y contaba con paciente simulador.

En Oftalmología una estación consistía en evaluar una imagen de ojo rojo y determinar el diagnóstico, la segunda estación consistía en determinar la agudeza visual y contaba con paciente simulador.

Universo muestral

Los estudiantes de 6° Año de Cirugía fueron 86; cada alumno fue evaluado mediante diez listas de cotejos correspondientes a diez estaciones de evaluación del ECOE.

2.2 Métodos y técnicas

La investigación se llevó a cabo utilizando las listas de cotejos de las diez estaciones del ECOE, cada una de las cuales constaba de la presencia de un profesor de las especialidades de Cirugía General, Neurocirugía, Oftalmología, Otorrinolaringología y Urología. Dos estaciones de Cirugía General utilizaron paciente simulador. En Neurocirugía una estación utilizó una placa de tomografía axial computarizada de un caso de traumatismo encéfalo craneano, la segunda estación utilizó la descripción de un caso de un paciente en coma al que se le debía dar una puntuación de acuerdo a la Escala de Glasgow. En Otorrinolaringología una estación utilizó un paciente simulador para realizar una prueba auditiva mediante un diapasón de Rinne, la otra utilizó una placa de tomografía axial computarizada de senos paranasales. En Urología una estación utilizó la presencia de un paciente simulador para cólico renal y la segunda utilizó una maqueta para colocar sonda de Foley. En Oftalmología, una estación utilizó una imagen de ojo rojo y la segunda utilizó un paciente simulador y una gráfica de Snellen para determinar agudeza visual.

De las diez estaciones, cinco fueron activas y contaban con la presencia de un profesor y un paciente simulador. El paciente simulador fue entrenado previamente por el docente de acuerdo al campo que requería ser evaluado.

Cinco estaciones fueron pasivas y contaron con la presencia de un profesor quien había dispuesto una imagen radiográfica o de tomografía axial computarizada o contaba con un maniquí o una maqueta.

Los estudiantes ingresaron a los ambientes de examen en número de diez, uno por cada estación de acuerdo a su número de orden. Antes de ingresar a cada cubículo, leen las indicaciones del instructivo en el que se expone brevemente el caso o el procedimiento que debe realizar el alumno), luego pasan al ambiente en el que se encuentra un profesor además de la imagen a observar y describir de parte del estudiante o también el maniquí o la maqueta en la que el estudiante debía realizar algún procedimiento. En otras estaciones se encontraba un profesor y un paciente simulador. Los alumnos realizaban la anamnesis o el examen físico de acuerdo a la patología planteada. Los estudiantes finalizaban la rotación respondiendo preguntas escritas en algunas estaciones. El tiempo de rotación por cada estación fue de cinco minutos incluyendo el tiempo para leer el instructivo; una

vez que completaban las rotaciones por las diez estaciones se retiraban e ingresaba el siguiente grupo de estudiantes.

Los profesores solamente observaron el comportamiento y rendimiento de los alumnos, y registraban en su lista de cotejos, de manera dicotómica: si se cumplían o no los distintos pasos considerados en las respectivas estaciones. Por ejemplo, si el alumno al ingresar saludaba o no al profesor y al paciente simulador presentes en la estación, igualmente, si el alumno se colocaba o no los guantes para realizar un procedimiento, si seguía un orden o era desorganizado, si establecía o no un diagnóstico, si sugería o no realizar exámenes complementarios, etc. La calificación era en base 100 para cada estación y de inmediato se convertía a la escala vigesimal. Posteriormente se obtuvieron los promedios de las calificaciones de las diez rotaciones por alumno siendo el promedio aritmético de éstas, la nota de ECOE. Esta calificación nos permitió determinar en escala vigesimal el nivel de rendimiento por competencias alcanzado por los alumnos.

Cada una de las estaciones se evaluó desde el punto de vista de la validez y confiabilidad de las listas de cotejos.

Se utilizó la división por mitades y la corrección de Spearman-Brown para determinar el grado de confiabilidad de las listas de cotejos y la validez predictiva que usa el coeficiente de correlación de Pearson para determinar la validez de las listas de cotejos. El análisis estadístico utilizó el programa SPSS versión 15.0

Se recogieron los datos asignados a cada uno de los ítems de todas las listas de cotejos con las que se evaluaron los 86 alumnos, los datos se recogieron en tablas y luego fueron sometidos a las pruebas estadísticas siguientes:

Para determinar la confiabilidad de las listas de cotejos se usó el método de división por mitades (Hernández 2006: 288-290) que usa la fórmula:

$$r_{ip} = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{(n \sum X^2 - (\sum X)^2)(n \sum Y^2 - (\sum Y)^2)}}$$

Dónde:

r_{ip} = Correlación de preguntas impares con preguntas pares

n = número de cuestionarios

x = suma de los valores de las preguntas impares

y = suma de los valores de las preguntas pares

y la corrección de SPEARMAN - BROWN, que usa la fórmula:

$$r_{xx} = \frac{2 r_{ip}}{1 + r_{ip}}$$

Dónde:

r_{xx} = Coeficiente de confiabilidad

Para la validez se usó la validez de criterio (Hernández 2006: 288-290) que establece la validez de un instrumento de medición al comparar sus resultados con los de algún criterio externo que pretende medir lo mismo fijado en el futuro, por lo que evaluamos la validez predictiva, que usa el coeficiente de correlación de Pearson, dado por:

$$r_{xx} = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{(n \sum X^2 - (\sum X)^2)(n \sum Y^2 - (\sum Y)^2)}}$$

Donde

r_{xx} = Coeficiente de validez

x = primera mitad de los valores totales

y = segunda mitad de los valores totales

El principio es: Si diferentes instrumentos o criterios miden el mismo concepto o variable, deben arrojar resultados similares. Luego de la valoración estadística, se establecieron los datos de validez y confiabilidad que se presentaron la Tabla 2.

III. RESULTADOS Y DISCUSIÓN

Las figuras 2 y 3 son ilustraciones del proceso de evaluación de los estudiantes de medicina del sexto ciclo mediante el ECOE.



Fig. 2. Los estudiantes leen el instructivo antes de ingresar a una estación del ECOE.



Fig. 3. El profesor observa y anota en su lista de cotejos si el estudiante cumple o no los pasos del procedimiento. En el presente caso, examen físico de abdomen.

En la tabla 1 se observa que el rendimiento por competencias quirúrgicas de los 86 estudiantes, según las notas obtenidas de la calificación de cada una de las listas de cotejos de las diez estaciones del ECOE, varía desde 8,1 hasta 18,4 puntos, siendo el promedio final de 16.2 puntos.

En la tabla 2, se observa que las listas de cotejos de evaluación, utilizando el método de división por mitades y la corrección de Sperman-Brown, tienen una confiabilidad que varía de 0,33 hasta 0,90 y en las estaciones 1, 7, 8 y 10 los instrumentos no son confiables. En cuanto a la validez de los diferentes instrumentos, determinada mediante el coeficiente de correlación de Pearson, estos varían desde 0.65 hasta 1.00.

Trejo Mejía et al. (2009) evaluaron con el formato ECOE a 74 internos de pregrado de la Facultad de Medicina de la Universidad Autónoma de México y encontraron que la confiabilidad de los exámenes fue 0,83, 0,76 y 0,75 en tres ocasiones.

Cohen et al. (1990) evaluaron con ECOE a 27 residentes de cirugía de segundo año, en la Universidad de Ontario en Toronto - Canadá; la confiabilidad total fue de 0.89 y la validez tuvo un rango entre 0.36 y 0.82.

Reznick et al, (1997) aplican un ECOE para habilidades técnicas a 48 residentes de la Facultad de Medicina de la Universidad de Manitoba, Canadá y encontraron una confiabilidad de 0.78 para las listas de cotejos y de 0.85 para la valoración global; la validez de constructo la determinaron con análisis de varianza y reveló un efecto significativo.

Las pruebas con las estimaciones de confiabilidad entre 0.0 a 0.50, son muy imprecisas, Las pruebas con índices de 0,50 a 0,80, son moderadamente fiables y las pruebas con índices de confiabilidad mayor a 0.80 pueden ser utilizadas con confianza para pruebas avanzadas como para la certificación (Reznick et al, 1997).

Grand'Maison et al (1992), en Quebec evaluando a 547 nuevos médicos de familia de cuatro escuelas de medicina en tres exámenes encontraron coeficiente de confiabilidad de 0,644, 0,723 y 0,736.

Wessel et al (2003), aplicaron un ECOE a 48 estudiantes de un programa de fisioterapia en Ontario, Canadá y encontraron una confiabilidad de 0.48 y una validez entre -0.14 y 0.33 con lo que concluyen que la evaluación que hicieron tenía una pobre consistencia interna y no predecía la performance clínica.

Tabla 1. Rendimiento por competencias quirúrgicas de los estudiantes por estación y según lista de cotejos

ESTACIÓN	Rendimiento de los alumnos mediante lista de cotejos																	PROMEDIO TOTAL	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	...		86
1.ADA	18	18	20	16	16	18	20	18	16	18	16	18	18	16	18	18		16	17.5
2.EFA	20	16	18	18	14	18	20	18	16	20	20	14	12	16	18	14		12	16.7
3.TAC	20	15	20	15	20	20	20	20	15	20	10	20	15	15	15	5		12	16.3
4.EG	20	20	20	20	20	20	20	0	20	20	20	20	20	12	20	20		15	18.1
5.AV	18	14	20	10	12	12	16	14	12	14	20	16	20	18	18	16		20	15.9
6.OR	8	0	2	10	11	6	8	2	12	0	20	7	16	7	18	10		0	8.1
7.SPN	18	14	18	20	18	20	20	19	16	20	18	20	18	18	20	18		18	18.4
8.PR	20	18	20	18	20	20	20	20	18	20	20	20	20	20	20	20		20	19.6
9.SF	14	12	8	6	16	12	18	12	14	16	12	16	14	14	16	10		10	12.9
10.CR	18	16	20	18	16	18	18	16	18	18	18	18	16	20	18	18		20	17.9
PROMEDIO TOTAL	17.4	14.3	17	15.1	16.3	16.4	18	13.9	15.7	16.6	17.4	16.9	16.9	15.6	18.1	14.9		14.3	16.2

ADA: anamnesis de abdomen agudo, EFA: examen físico de abdomen, TAC: tomografía axial computarizada, EG: escala de Glasgow, AV: agudeza visual, OR: ojo rojo, SPN: senos para nasales, PR: prueba de Rinne, SF: sonda de Foley, CR: cólico renal.

Tabla 2. Validación de instrumentos de medición

LISTAS DE COTEJOS	CONFIABILIDAD	VALIDEZ
1.Ananesis de Dolor Abdominal	0	0,73
2.Examen Físico Abdominal	0,76	0,90
3.Tomografía Axial Computarizada	0,50	0,86
4.Escala de GLASGOW	0,90	1,00
5.Agudeza Visual	0,57	0,93
6.Ojo Rojo	0,80	0,96
7.Senos Para Nasaes	0	0,94
8.Prueba de Rinne	0	0,65
9.Sonda de Foley	0,33	0,94
10.Cólico Renal	0	0,76

En nuestro estudio se apreció que la confiabilidad alcanzó niveles de 0.33 a 0.90 en seis estaciones y cuatro no tuvieron confiabilidad; las estaciones no confiables, fueron una de cirugía general la que se refiere a anamnesis del dolor abdominal; dos de otorrinolaringología, la prueba de Rinne y la prueba de senos para nasaes y una de urología, la referida a la prueba de cólico renal. La validez varió desde 0.65 hasta 1.0 (Tabla 2).

El proceso de evaluación mediante el ECOE en el Departamento de Cirugía, requiere ser sometido a ajustes, para que los resultados puedan ser consistentes y repetibles.

Al comparar nuestros resultados con los de otros autores, nos damos cuenta que existen instituciones que aplican ECOE y reportan diferentes niveles de confiabilidad y validez de sus pruebas.

Estos datos nos permiten inferir que en nuestra realidad como en muchas otras, la enseñanza aprendizaje del médico privilegia el conocimiento y es muy poca o casi nula la formación en aspectos educativos. Mientras más experiencia y capacitación de los profesores, es mejor el diseño de las pruebas y mejores los resultados de confiabilidad y validez.

El proceso de formación de los médicos es largo y complejo, y requiere que el estudiante asimile conocimientos científicos, desarrolle habilidades y se forme o adquiera actitudes y valores para su desempeño profesional.

La evaluación de las competencias es parte del proceso de enseñanza aprendizaje.

La sociedad espera que los médicos que tengan que desenvolverse atendiendo a la comunidad, estén formados en competencias que involucren conocimientos, destrezas y valores.

V. CONCLUSIONES

El rendimiento por competencias quirúrgicas de los 86 estudiantes de medicina del sexto año de la Universidad Nacional de Trujillo, según los calificativos obtenidos de cada una de las listas de cotejos de las diez estaciones del ECOE (Examen Clínico Objetivo Estructurado) fue muy satisfactorio. El calificativo promedio fue de 16.2 puntos.

La evaluación de las 10 listas de cotejos de las pruebas ECOE aplicado a los estudiantes, demuestra que las diez cuentan con confiabilidad satisfactoria. La confiabilidad alcanzó niveles de 0.33 a 0.90 en seis estaciones y cuatro no tuvieron confiabilidad.

En cuanto a la validez, cuatro estaciones cuentan con listas de cotejos válidas para evaluaciones parciales y seis estaciones tienen listas de cotejos imprecisas. La validez de los diferentes instrumentos, determinada mediante el coeficiente de correlación de Pearson, estos varían desde 0.65 hasta 1.00.

VI. RECOMENDACIONES

Se requiere la realización de trabajo en equipo de los profesores para hacer ajustes en la elaboración de las estaciones de ECOE de acuerdo al perfil epidemiológico de las patologías.

Es necesaria una capacitación permanente de los profesores de Cirugía, en cuanto a técnicas de validez y confiabilidad de las pruebas, a cargo de profesores expertos del Programa de Educación.

Mediante trabajo multidisciplinario, se deben revisar las estaciones de ECOE para mejorar la confiabilidad.

VII. REFERENCIAS BIBLIOGRÁFICAS

- AIKEN, L. R. 2003. **Tests Psicológicos y Evaluación** (11a ed.). México: PEARSON Education. (<http://books.google.com.pe/books?id=2LvyL8JEDmQC&printsec=frontcover&hl=es#v=onepageq&f=false>, consultado el 09 de enero, 2012).
- ANDRADE CÁZARES, R. A., HERNÁNDEZ, S. C. 2010. **El enfoque de competencias y el currículum del bachillerato en México**. (<http://www.redalyc.org/articulo.oa?id=77315079023>, consultado el 10 de abril, 2011).
- BAZÁN GUZMÁN, J. 2000. **Evaluación psicométrica de las preguntas y pruebas – Crecer 96**. (http://www2.minedu.gob.pe/umc/admin/images/menanexos/menanexos_206.pdf, consultado el 04 de marzo, 2012).
- BENEITONE, P., ESQUETINI, C., GONZÁLES, J., MARTY, M., SIUFI, G., WAGENAAR R. 2007. **Proyecto Tuning América Latina - Informe Final: Reflexiones y perspectivas de la Educación Superior en América Latina**. Universidad de Deusto, Universidad de Groningen. Donostia – San Sebastian. Guipuzcoa, España. (http://tuning.unideusto.org/tuningal/index.php?option=com_docman&Itemid=191&task=view_category&catid=22&order=dmdate_published&asc=DESC, consultado el 10 de junio, 2011).
- COHEN R., REZNICK R. K., TAYLOR B. R., PROVAN J., ROTHMAN A. 1990. **Reliability and validity of the objective structured clinical examination in assessing surgical residents**. 1990. Amer. J. Surg., Vol. 160:302-305.
- CUSCHIERI, A., GLEESON, F. A., HARDEN, M., WOOD, R. 1979. **A new approach to a final examination in surgery**. Ann. Royal Coll. Surg. England, Vol. 61: 400-405.
- DECLARACIÓN DE BOLONIA. 1999. **Espacio europeo de enseñanza superior. Declaración conjunta de los ministros europeos de enseñanza**. (<http://eees.umh.es/contenidos/Documentos/DeclaracionBolonia.pdf>, consultado el 20 de febrero, 2011).
- DOWING, S. M. 2004. **Reliability on the reproducibility of assessment data**. Med. Educ., Vol. 38:1006-1012.
http://medicina.udd.cl/ode/files/2010/07/Downing_reliability_2490.pdf, consultado el 12 de enero, 2012).

- ESCUELA ANDALUZA DE SALUD PÚBLICA. 2008. **Diseño de instrumentos de evaluación de competencias y desarrollo profesional. Curso de Gestión de la formación continuada en profesiones sanitarias Granada.** (<http://www.slideshare.net/jimenezpemettp/portfolios-las-competencias-en-el-desarrollo-profesional-presentation>, consultado el 10 de enero, 2011).
- GONZÁLEZ, J., WAGENAAR, R. 2003. **Tuning Educational Structures in Europe. Informe Final. Fase Uno.** Bilbao, España: RGM, S.A. (http://www.relint.deusto.es/TUNINGProject/spanish/doc_fase1/Tuning%20Educational.pdf, consultado el 10 de junio, 2011).
- GRAND'MAISON, P., LESCOPE, J., RAINSBERRY, P., BRAILOVSKY, C. A. 1992. **Large-scale use of an objective, structured clinical examination for licensing family physicians.** Can. Med. Assoc. J., Vol. 146 (10): 1735-1740.

(<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1488712/pdf/cmaj00299-0057.pdf>, consultado el 19 de abril, 2012).
- HARDEN, R., STEVENSON, M., WILSON DOWNIE, W., WILSON, G. M. 1975. **Assessment of Clinical Competence using Objective Structured Examination.** British Medical Journal, Vol. 1: 447-451.

(<http://www.bmj.com/content/1/5955/447.pdf%2Bhtml>, consultado el 22 de febrero, 2011).
- HERNÁNDEZ R., FERNÁNDEZ C. y BAPTISTA P. 2006. **Metodología de la Investigación.** (4ta ed.). México: McGraw Hill.
- JONES, R., HIGGS, R., ANGELIS, C., PRIDEAUX, D. 2001. **Changing face of medical curricula.** The Lancet, Vol. 357: 699-703.
- JOORABCHI B. 1991. **Objective structured clinical examination in a pediatric residency program.** Am. J. Dis. Child., Vol. 145: 757-762.
- LOZANO SÁNCHEZ, R. 2010. **¿Qué son las competencias?**

(<http://www.slideshare.net/Rodegarius/qu-son-las-competencias-5026220>, consultado el 10 de enero, 2011).
- UNIVERSIDAD NACIONAL DE TRUJILLO. FACULTAD DE MEDICINA. 2011. **Sílabo de Cirugía II. 6° Año de Medicina. Trujillo – Perú**
- MILLER G.E. 1990. **The assessment of clinical skills/competence/performance.** Acad. Med., Vol. 65: 563-67.
- PETRUSA E. R., BLACKWELL T. A. & AINSWORTH M. A. 1990. **Reliability and validity of an objective structured clinical examination for assessing the clinical performance of residents.** Arch. Intern. Med., Vol. 150: 573-7.
- RETA, A.M., LÓPEZ, M.J., MONTBRUN, M., ORTIZ, A., VARGAS, A.L. 2006. **Competencias médicas y su evaluación al egreso de la carrera de medicina en la Universidad Nacional de Cuyo (Argentina).** Educ. Med., Vol. 9: 75-83.
(http://scielo.isciii.es/scielo.php?pid=S1575-18132006000200006&script=sci_arttext, consultado el 28 de marzo, 2011).
- REZNICK, R., REGEHR, G., MACRAE, H., MARTIN, J., MCCULLOCH, W. 1997. **Testing Technical skill via an innovative “Bench station” Examination.** Amer. J. Surg., Vol. 173:226-230.
(http://www.umanitoba.ca/faculties/medicine/units/surgery/general_program/media/Bench.pdf, consultado el 12 de marzo, 2012).
- ROSS M., CARROLL G., KNIGHT J., CHAMBERLAIN M., FOTHERGILL-BOURBONNAIS F. & LINTON J. 1988. **Using the OSCE to measure clinical skills performance in nursing.** J. Adv. Nurs., Vol. 13: 45-46. (<http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2648.1988.tb01390.x/abstract>, consultado el 04 de junio, 2011).

- SÁNCHEZ MENDIOLA, M. 2012. **La calidad del proceso de evaluación para la Certificación del médico especialista.** (http://www.conacem.org.mx/files/boletin_calidad.pdf, Consultado el 10 de noviembre, 2012).
- SLOAN, D. A., DONNELLY, M. B., JOHNSON, S. B., SCHWARTZ, R.W., STRODEL, W. E. 1993. **Use of an Objective Structured Clinical Examination (OSCE) to measure improvement in clinical competence during the surgical internship.** *Surgery*, Vol. 114: 343-50.
- SLOAN, D. A., DONNELLY, M. B., SCHWARTZ, R., STRODELL, W. 1995. **The Objective Structured Clinical Examination. The New Gold Standard for Evaluating Postgraduate.** *Clinical Performance Ann. Surg.* Vol. 222: 735-742.
- SLOAN, D., DONNELLY, M., SCHWARTZ, R., FELTS, J., BLUE, A., STRODEL, W. 1996. *Journal Of Surgical Research*; 63: 225-230.
- SOCIEDAD ESPAÑOLA DE MÉDICOS DE ATENCIÓN PRIMARIA (SEMERGEN) .2009. **Prueba ECOE.** (<http://multimedia.diariomedico.com/videos/prueba-ecoe>, consultado el 10 de abril, 2011).
- TREJO MEJÍA, J. A., PEÑA BALDERAS, J., MARTÍNEZ GONZÁLEZ, A. , SÁNCHEZ MENDIOLA M. 2009. **El desarrollo de nuevas competencias de ECOE para la evaluación de la competencia clínica en el internado médico.** (<http://www.facmed.unam.mx/sem/jem2010/Carteles/Viernes/EV13.pdf>, consultado el 15 de marzo, 2012).
- TUNING AMÉRICA LATINA. 2004-2008. **Competencias Específicas de Medicina.** (<http://tuning.unideusto.org/tuningal/index.php?option=content&task=view&id=229&Itemid=258>, consultado el 20 de febrero, 2011).
- UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO. FACULTAD DE MEDICINA, s/f. **Evaluación de competencias profesionales en la carrera de médico cirujano.** (<http://www.facmed.unam.mx/eventos/competencia/001.pdf>, consultado el 03 de marzo del 2011).
- UNIVERSIDAD NACIONAL DE TRUJILLO. FACULTAD DE MEDICINA. 2011. **Sílabo de Cirugía II. 6° Año de Medicina.** Trujillo – Perú.
- WASS V., VAN DER VLEUTEN C., SHATZER J., JONES, R. 2001. **Assessment of clinical competence.** *The Lancet*, Vol. 357: 945-949.
(<http://download.thelancet.com/pdfs/journals/lancet/PIIS0140673600042215.pdf?id=410a13c7e56fa01:-74ba94ef:13e7817f8a5:-11041367817236766>), consultado el 03 de mayo, 2011).
- WESSEL, J., WILLIAMS, R., FINCH, E. GÉMUS, M. 2003. **Reliability and validity of an Objective structured clinical examination for physical therapy students.** *J. Allied Health*, Vol. 32: 266-269.

ANEXO

ECOE: ESTACIÓN DE CIRUGÍA GENERAL ANAMNESIS DE DOLOR ABDOMINAL

Examinador:

CALIFICACIÓN

Alumno:

Fecha:

INSTRUCCIONES PARA EL OBSERVADOR

Limítese a observar las acciones del alumno: interrogación al paciente sobre semiología del dolor, sin sugerir ni corregir ningún comentario.

Anote si sigue la secuencia sugerida en la lista y califique.

LISTA DE PREGUNTAS

HABILIDAD TECNICA	100 PUNTOS	SI	NO
1. Saluda y explica al paciente sobre La entrevista a realizar.	(10 p)	<input type="checkbox"/>	<input type="checkbox"/>
2. Tiempo de duración del dolor.	(10 p)	<input type="checkbox"/>	<input type="checkbox"/>
3. Localización e irradiación del dolor	(10 p)	<input type="checkbox"/>	<input type="checkbox"/>
4. Calidad del dolor.	(10 p)	<input type="checkbox"/>	<input type="checkbox"/>
5. Intensidad del dolor.	(10 p)	<input type="checkbox"/>	<input type="checkbox"/>
6. Factores desencadenantes.	(10 p)	<input type="checkbox"/>	<input type="checkbox"/>
7. Factores que alivian o agravan.	(10 p)	<input type="checkbox"/>	<input type="checkbox"/>
8. Concomitantes: náuseas o vómitos	(10 p)	<input type="checkbox"/>	<input type="checkbox"/>
9. Variación del dolor con alimentos	(10 p)	<input type="checkbox"/>	<input type="checkbox"/>
10. Evalúa antecedentes de ictericia	(10 p)	<input type="checkbox"/>	<input type="checkbox"/>