



Original Article

## Evaluación objetiva de competencias quirúrgicas en cirugía mínimamente invasiva con una colección de pruebas simples

Objective evaluation of surgical competency for minimally invasive surgery with a collection of simple tests

Eliana Maria Gonzalez-Neira<sup>1</sup>, Claudia Patricia Jimenez-Mendoza<sup>1</sup>, Daniel R Suarez<sup>1</sup>, Saul Rugeles-Quintero<sup>2</sup>

<sup>1</sup> Departamento de Ingeniería Industrial, Facultad de Ingeniería, Pontificia Universidad Javeriana, Bogotá, Colombia.

<sup>2</sup> Departamento de Cirugía, Pontificia Universidad Javeriana, Hospital Universitario San Ignacio, Bogotá, Colombia.

Gonzalez-Neira EM, Jimenez-Mendoza CP, Suarez DR, Rugeles-Quintero S. Objective evaluation of surgical competency for minimally invasive surgery with a collection of simple tests. *Colomb Med (Cali)*. 2016; 47(1):45-50.

© 2016 Universidad del Valle. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Attribution License, que permite el uso ilimitado, distribución y reproducción en cualquier medio, siempre que el autor original y la fuente se acreditan.

### Historia:

Recibido: 26 septiembre 2014  
Revisado: 16 diciembre 2015  
Aceptado: 12 enero 2016

### Palabras clave:

Entrenamiento basado en competencias, Desempeño en sala de cirugía, Evaluación de habilidades para laparoscopia, Simulación

### Keywords:

Proficiency-based training, laparoscopic skill assessment, simulation

### Resumen

**Objetivo:** Este estudio pretende determinar si una colección de 16 pruebas en un simulador físico puede discriminar y evaluar objetivamente el nivel de competencia de practicantes de cirugía laparoscópica (novato, resistente y experto).

**Métodos:** Se realizó un diseño de experimentos con tres grupos de estudio (novatos, residentes y expertos) para probar el poder de discriminación de 16 pruebas simples. Un ANOVA y un test de Student Newman-Keuls (SNK) fueron empleados para analizar los resultados de cada prueba con el fin de determinar cuáles de ellas podían discriminar los niveles de competencia de los participantes.

**Resultados:** Cuatro de las 16 pruebas evaluadas lograron discriminar los tres niveles de competencia, y 15 pruebas lograron discriminar al menos dos de los tres grupos ( $\alpha = 0.05$ ). Adicionalmente, otras dos pruebas diferenciaron los novatos de los residentes, y otras siete pruebas diferenciaron los residentes de los expertos.

**Conclusión:** El nivel de competencia de un practicante de cirugía mínimamente invasiva puede ser determinado mediante una colección de pruebas básicas en un simulador quirúrgico físico. El diseño de pruebas que discriminen los tres niveles de competencia y reduzcan el número de pruebas de la colección son posibles temas para nuevos trabajos.

### Abstract

**Objective:** This study aims at determining if a collection of 16 motor tests on a physical simulator can objectively discriminate and evaluate practitioners' competency level, i.e. novice, resident, and expert.

**Methods:** An experimental design with three study groups (novice, resident, and expert) was developed to test the evaluation power of each of the 16 simple tests. An ANOVA and a Student Newman-Keuls (SNK) test were used to analyze results of each test to determine which of them can discriminate participants' competency level.

**Results:** Four of the 16 tests used discriminated all of the three competency levels and 15 discriminated at least two of the three groups ( $\alpha = 0.05$ ). Moreover, other two tests differentiate beginners' level from intermediate, and other seven tests differentiate intermediate level from expert.

**Conclusion:** The competency level of a practitioner of minimally invasive surgery can be evaluated by a specific collection of basic tests in a physical surgical simulator. Reduction of the number of tests needed to discriminate the competency level of surgeons can be the aim of future research.

## Introduction

La cirugía mínimamente invasiva (CMI) permite llevar a cabo procedimientos quirúrgicos sin necesidad de realizar una gran incisión, disminuyendo las lesiones en el paciente, el riesgo de hemorragias, y el tiempo de recuperación postoperatoria<sup>1</sup>. Sin embargo, regularmente se reporta un limitado campo visual y alteración de la percepción de profundidad y fuerza<sup>2</sup>. Esto demanda del cirujano competencias que se adquieren y perfeccionan únicamente con la práctica de la técnica, ya que no son intuitivas, ni comunes durante actividades cotidianas o la cirugía abierta convencional<sup>3</sup>.

Actualmente, la comunidad médica centra gran atención en investigaciones para el desarrollo de técnicas de adquisición, entrenamiento, y evaluación de las habilidades que requiere la CMI<sup>4-7</sup>. Las referencias regionales en el tema, para el conocimiento de los autores de este estudio, es casi inexistente<sup>8</sup>. Las evaluaciones y acreditación final de las competencias de un cirujano están enmarcadas en un método de aprendizaje del tipo «maestro-discípulo», y continúan siendo una parte fundamental en la instrucción de los practicantes de CMI. Específicamente, la evaluación de las competencias del practicante se basan generalmente en el recuerdo del rendimiento del practicante por parte del maestro y, por lo tanto, están sujetas a error<sup>3</sup>. Las metodologías de evaluación incluyen el método «Objective Structured Assessment of Technical Skills» (OSATS)<sup>9</sup> y herramientas de observación para evaluación de habilidades quirúrgicas. Sin embargo, la evaluación de competencias técnicas usando métodos observacionales no es confiable y válido para todos los niveles de competencia<sup>10</sup>.

Varios simuladores comercialmente disponibles son utilizados para aprender y entrenar las habilidades básicas en CIM<sup>11</sup>. La práctica con simuladores de realidad virtual permite al médico experimentar condiciones parecidas a las reales, pero sin el riesgo y el costo de una cirugía real. Por otra parte, los simuladores físicos (o cajas de entrenamiento) permiten trabajar en un espacio tridimensional con una percepción táctil y de profundidad más acertada, y a costo aún más bajo que el de un simulador de realidad virtual<sup>12</sup>. Sin embargo, estos simuladores han sido criticados por ser poco realistas y alejados de dar una evaluación objetiva de los procedimientos laparoscópicos reales<sup>13</sup>.

En general, el proceso de aprendizaje en CMI debe incluir evaluaciones objetivas que permitan conocer y acreditar la evolución y nivel de experticia de las competencias de los practicantes. La discriminación de los diferentes niveles de competencia en CMI es entonces el primer requerimiento fundamental de un método de evaluación para practicantes. Así, el objetivo de este estudio es determinar si una colección de pruebas sencillas en un simulador físico puede discriminar el nivel de competencia de un practicante en CMI. Se asume que la(s) prueba(s) de evaluación ideal(es) debe(n), al menos, discriminar entre tres niveles de competencia: Novato, Residente y Experto.

## Materiales y Métodos

La metodología en este estudio incluye tres partes secuenciales: definición de la colección de pruebas, un experimento y el análisis de los resultados.

### Definición de la colección de pruebas

Un equipo interdisciplinario que incluyó un cirujano experto, estudió y modificó varias pruebas de la literatura. Un total de 16 pruebas fueron implementadas para este estudio. El propósito de cada prueba fue evaluar una de cuatro habilidades fundamentales en CIM: desplazamiento (De), corte (C), disección (Di), y sutura (S). Cada habilidad fundamental fue evaluada por cuatro pruebas. Los criterios de objetividad, economía y simplicidad fueron tomados en cuenta para la selección de las pruebas. Las pruebas seleccionadas se describen a continuación:

#### Actividades de Desplazamiento (Fig. 1)

De1-»Cilindros»: Ordenar unos cilindros de izquierda a derecha y del más pequeño al más grande en las dos primeras filas de un tablero.

De2-»Cubos»: Esta actividad requiere armar dos cubos o cajas a partir de cuatro fichas y una base.

De3-»Torre»: Armar una torre con cuatro cubos y ubicar seis fichas en determinado lugar de la tabla de acuerdo a su geometría.

De4-»Palillos»: Atravesar ocho palillos a través de dos de los orificios de un tubo rígido.

#### Actividades de corte (Fig. 2)

C1-»Aluminio»: Cortar tres cuadrados de 5 cm de lado en papel aluminio siguiendo unas líneas guías.

C2-»Figura»: Cortar un rectángulo dibujado sobre un papel.

C3-»Medio globo»: Cortar un círculo dibujado sobre medio globo elástico.

C4-»Malla»: Cortar los hilos de color blanco ubicados en una malla de hilos de colores, tratando de no tocar los demás.

#### Actividades de Disección (Fig. 3)

Di1-»Plástico»: Quitar el plástico que envuelve una figura de plastilina sin causarle daños a la misma.

Di2-»Operando»: Sacar tres esferas de plásticas de un globo, después de realizar un corte indicado por una línea guía.

Di3-»Canica»: Quitar el papel aluminio que cubre una canica causando el menor daño posible a la envoltura.

Di4-»Espuma»: Remover el cubo de espuma adherido a la parte inferior de un objeto procurando que la espuma sufra el menor daño posible.

#### Actividades de Sutura (Fig. 4)

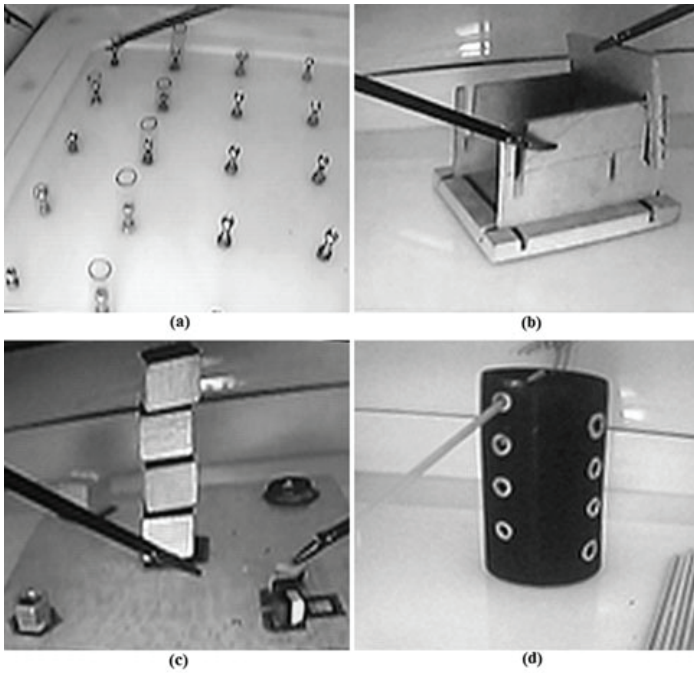
S1-»Plastilina»: Anudar tres barras de plastilina pasando un hilo alrededor de ellas, de forma que queden unidas.

S2-»Campanas»: Amarrar tres campanas a un palo después de atravesar con aguja e hilo sus orificios superiores.

S3-»Ramo»: Anudar el extremo inferior de cuatro globos, de acuerdo a un punto guía.

S4-»Collar»: Hacer un collar con seis fichas, atravesando con aguja e hilo los orificios correspondientes y luego anudándolas.

Cada una de las pruebas se evaluaron de acuerdo a unos indicadores característicos para cada actividad: En los ejercicios de desplazamiento se evaluó el tiempo, las veces que fueron soltados los objetos, la precisión en la ubicación de objetos y el número de objetos ubicados correctamente. En los ejercicios de corte se evaluó el tiempo, los hilos cortados y tocados, la longitud de corte y la precisión del mismo. En los ejercicios de sutura se evaluó el tiempo, los nudos (si se soltaban o no), el número de

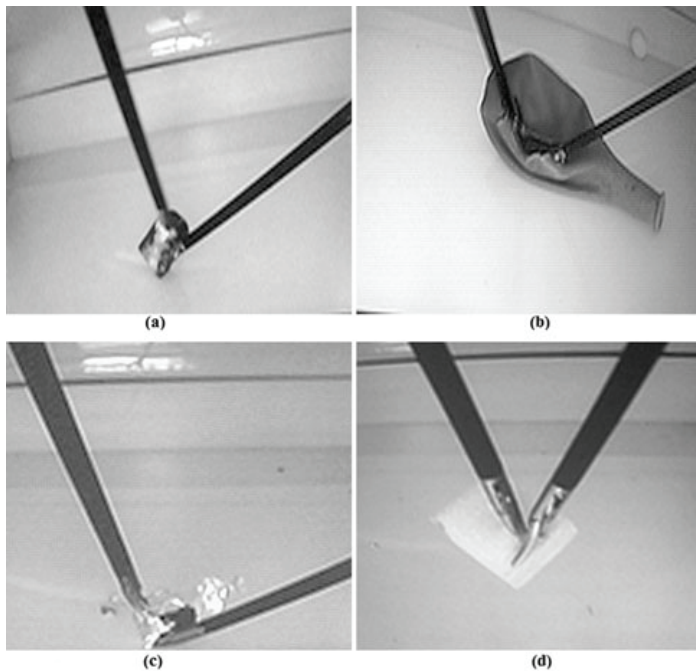


**Figura 1.** Pruebas de desplazamiento: (a) De1-"Cilindros": ordenar 10 cilindros del más pequeño al más grande en un tablero, (b) De2-"Cubos": construir dos cajas a partir de cinco piezas de madera, (c) De3-"Torre": construir una torre hecha de cuatro cubos y localizar cuatro objetos alrededor de la torre, y (d) De4-"Palillos": atravesar orificios predefinidos en un cilindro con ocho palillos.

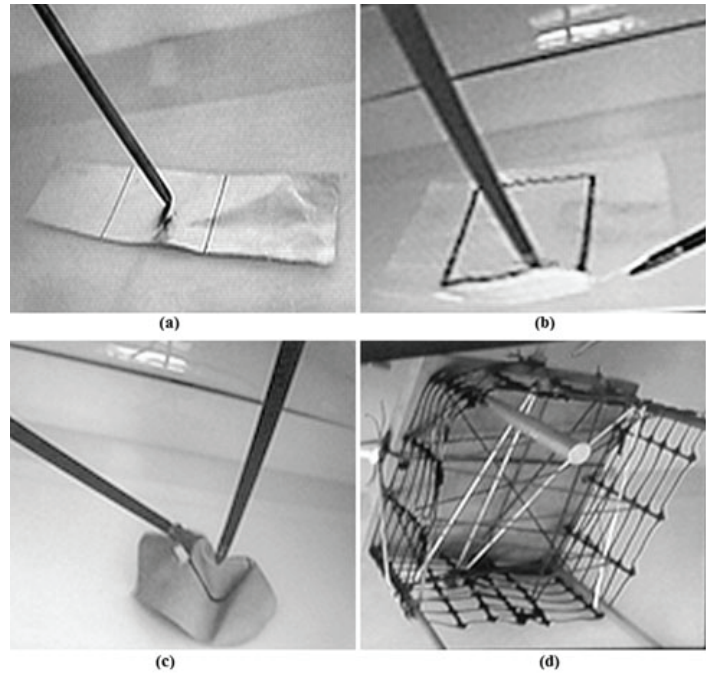
objetos atravesados y su precisión. Por último, en los ejercicios de disección se evaluó el tiempo, el número de daños a los objetos manipulados (huecos o rasgaduras), el estado de los materiales y el número de esferas sacadas del globo.

### Experimento

La hipótesis de este estudio es si una colección de pruebas simples



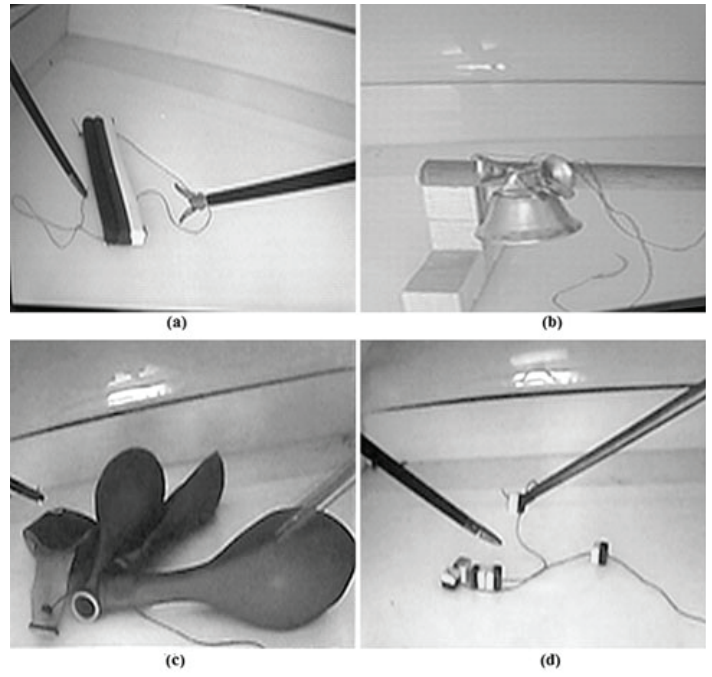
**Figura 3.** Pruebas de disección: (a) Di1-"Plástico": quitar el plástico que envuelve una figura de plastilina sin causarle daños a la misma, (b) Di2-"Operando": sacar tres esferas de plásticas de un globo, después de realizar un corte indicado por una línea guía, (c) Di3-"Canica": quitar el papel aluminio que cubre una canica causando el menor daño posible a la envoltura, y (d) Di4-"Espuma": remover el cubo de espuma adherido a la parte inferior de un objeto procurando que la espuma sufra el menor daño posible.



**Figura 2.** Pruebas de corte: (a) C1-"Aluminio": cortar tres cuadros de aluminio siguiendo las líneas guía predefinidas, (b) C2-"Figura": cortar un rectángulo siguiendo una línea pre-establecida en una hoja de papel, (c) C3-"Medio globo": cortar un círculo demarcado en medio globo elástico, y (d) C4-"Malla": Cortar los hilos de color blanco ubicados en una malla de hilos de colores.

en una caja de entrenamiento es capaz de discriminar a los practicantes de CMI en una de tres categorías (novatos, residentes y cirujanos expertos). Un diseño experimental se utilizó para probar la hipótesis.

Tres grupos de practicantes de CMI fueron voluntarios en este estudio. El primer grupo, Novatos (A), se compuso de seis



**Figura 4.** Pruebas de sutura: (a) S1-"Plastilina": anudar tres barras de plastilina pasando un hilo alrededor de ellas, de forma que queden unidas, (b) S2-"Campanas": amarrar tres campanas a un palo después de atravesar con aguja e hilo sus orificios superiores, (c) S3-"Ramo": anudar el extremo inferior de cuatro globos, de acuerdo a un punto guía, y (d) S4-"Collar": hacer un collar con seis fichas, atravesando con aguja e hilo los orificios correspondientes y luego anudándolas.

estudiantes de nuestra Facultad de Medicina, quienes no tenían ninguna experiencia previa en cirugía. El segundo grupo, Residentes (R) consistió en seis personas que se encontraban en sus años finales de residencia en cirugía. El tercer y último grupo, Expertos (E), fueron seis cirujanos expertos reconocidos en procedimientos de CMI. Todos los voluntarios (18 en total) fueron reclutados en la Facultad de Medicina de la Pontificia Universidad Javeriana (Bogotá, Colombia) y firmaron un consentimiento informado que indicaba el riesgo de participar en el estudio (es decir, un riesgo menor), el derecho a dejar el estudio en cualquier momento, y el tratamiento privado de información personal y datos recolectados.

En un orden aleatorio, el total de los 18 voluntarios realizó las pruebas simples descritas. La ejecución del orden de las pruebas fue también aleatoria. Antes de llevar a cabo cada prueba, los participantes vieron un video en el cual un cirujano describía las pruebas a través de una demostración y explicaba cómo estas eran evaluadas. Se seleccionó un único evaluador para cada tipo de habilidad fundamental (desplazamiento, corte, sutura y disección). La evaluación de cada prueba fue realizada inmediatamente después de completada la misma. El experimento completo fue llevado a cabo en un total de cinco días.

### Análisis estadístico

Para facilitar el análisis de los resultados, todos los indicadores evaluados fueron normalizados de acuerdo con el desempeño de los voluntarios, así:  $R_{tin} = BS_{in} - S_{tin} / BS_{in} - WS_{in}$

donde  $R_{tin}$  es el desempeño relativo del voluntario  $t$  en el indicador de evaluación  $n$  en la prueba  $i$ ,  $BS_n$  y  $WS_n$  son el mejor y peor puntaje para el indicador de evaluación  $n$  obtenido por cualquiera de los participantes de esa prueba, y  $S_{tin}$  es la calificación obtenida por el participante  $t$  en la correspondiente prueba e indicador de evaluación.

Adicionalmente, un único índice de desempeño por prueba

( $\bar{R}_{ti}$ ) fue calculado por participante al promediar todos los índices relativos de una misma prueba  $i$ .

Se utilizó un diseño completamente aleatorizado de un factor (nivel de competencia o experticia) para todos los experimentos con tres tratamientos (novato, residente y experto). Por tanto, se realizó un análisis de varianza ANOVA para cada una de las 16 pruebas. Se seleccionó una significancia de 0.05 para probar las hipótesis de si existe un efecto significativo del nivel de competencia o grupo al que el participante pertenece en los resultados de cada prueba ( $\bar{R}_{ti}$ ). Debido a que los ANOVA dieron resultados significativos, se realizó una prueba SNK para la comparación de tratamientos, de tal manera que se pudiera determinar cuáles grupos se podían discriminar significativamente con cada prueba simple. Para hacer estos análisis estadísticos se utilizó Excel 2010 (Microsoft, USA).

## Resultados

Los valores promedio de los índices de desempeño ( $\bar{R}_{ti}$ ) y su desviación estándar para cada grupo de estudio y actividad se presentan en la Figura 5. El grupo Experto siempre obtuvo un mejor índice de desempeño en las pruebas, excepto en una prueba de disección (Di1-»Plástico»), y el grupo Intermedio siempre tuvo un mejor índice de desempeño que el Grupo Novato, excepto en una prueba de desplazamiento (De3-Torre).

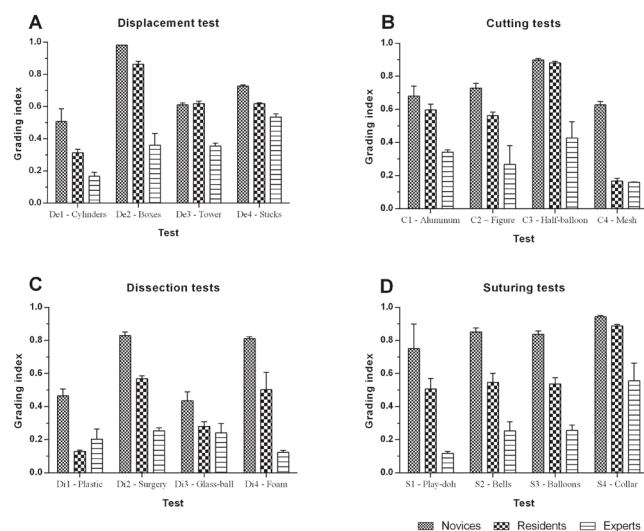
Los resultados del ANOVA apuntan a que todas las pruebas simples fueron capaces de discriminar al menos uno de los grupos de estudio, excepto la prueba Di3-»Canica» ( $p > 0.05$ ). La prueba de comparación de tratamientos SNK permitió concluir que solo cuatro pruebas fueron capaces de discriminar los tres grupos de estudio, esas fueron: Di2-»Operando», Di4-»Espuma» (ambas pruebas de disección), S2-»Campanas», and S3-»Ramo» (ambas pruebas de sutura). Ninguna prueba de desplazamiento ni corte pudo discriminar los tres grupos de estudio. Los resultados completos del el método SNK se pueden ver en la Tabla 1.

Finalmente, varias pruebas discriminaron exclusivamente los grupos Novatos y Residentes (C4-»Malla», and Di1-»Plástico») y los grupos Residentes y Expertos (De2-»Cubos, De3-»Torre», C1-»Aluminio», C2-»Figura», C3-»Medio globo», S1-»Plastilina» and S4-»Collar»).

## Discusión

Las ventajas de las pruebas en simuladores físicos o «box trainers» incluyen la retroalimentación táctil y la percepción de profundidad, ambos reales en los simuladores físicos. Sin embargo, la ausencia de pruebas de evaluación objetivas<sup>13</sup>, y la baja fidelidad de las actividades representadas son desventajas reconocidas en los simuladores físicos que deben ser superadas en el futuro<sup>7</sup>. El objetivo de este estudio es determinar si una colección de pruebas sencillas en un simulador físico, aun con las desventajas conocidas, puede discriminar el nivel de competencia de un practicante en CMI. Se asume que la discriminación entre niveles de competencia es un paso clave y necesario hacia el diseño de una evaluación objetiva de los practicantes de CMI, y cuya inexistencia es una desventaja mayor de los simuladores físicos.

Este estudio propuso inicialmente 16 pruebas en un simulador



**Figura 5.** Gráfico dinámico para la media y desviación estándar de los valores de la variable  $R_{ti}$  para todas las pruebas y grupos. Las barras en colores grises representan la media del índice de desempeño relativo de cada grupo (Novatos, Residentes y Expertos) en cada prueba. Las líneas verticales negras en la parte superior de las barras representan la desviación estándar del índice relativo de desempeño.

**Tabla 1.** Resultados de la prueba de comparación de tratamientos Student Newman Keuls (SNK) ( $\alpha=0.05$ )

Descripción	Agrupamiento dado por la prueba SNK (con 5% de significancia)*	Promedio de Rti de la colección de pruebas											
		S1	S2	S3	S4	C1	C2	C3	C4	Di1	Di2	De3	Di4
Pruebas que discriminaron los tres grupos	A Novatos		0.852	0.838								0.830	0.810
	B Residentes		0.546	0.537								0.569	0.503
	C Expertos		0.254	0.256								0.254	0.124
Pruebas que discriminaron dos grupos en los que Novatos y Residentes no son significativamente diferentes entre sí, y Expertos tienen un desempeño significativamente diferente a ellos	A Novatos	0.751			0.944	0.680	0.728	0.899				0.982	0.610
	A Residentes	0.507			0.889	0.596	0.562	0.881				0.863	0.619
	B Expertos	0.117			0.556	0.341	0.267	0.427				0.361	0.355
Novatos están en un grupo diferente a Residentes y Expertos (los cuales están en agrupados)	A Novatos										0.627	0.465	
	B Residentes										0.166	0.130	
	Expertos										0.159	0.204	
Los Novatos se diferencian claramente de los Expertos pero el desempeño de los Residentes no es significativamente diferente ni de los Novatos ni de los Expertos	A Novatos											0.507	
	A Residentes											0.314	
	B Expertos											0.168	
Las pruebas no pudieron discriminar a ninguno de los grupos, es decir, los Novatos, Residentes y Expertos no son significativamente diferentes entre sí	Novatos												0.436
	A Residentes												0.280
	Expertos												0.241

\*con 5% of confianza

S: sutura

C: corte

D:dissección

físico que involucraron actividades de desplazamiento, corte, disección y sutura. Sin embargo, no todas las pruebas en el simulador discriminaron el nivel de competencia de los participantes. Únicamente cuatro de las pruebas lograron discriminar los tres grupos de estudio (Expertos, Residentes, y Novatos). Por tanto, estas cuatro pruebas (Di2-»Operando», Di4-»Espuma», S2-»Campanas» y S3-»Ramo») son ideales para formar parte de una evaluación de competencias en CMI. Resultados similares han sido reportados. Un estudio<sup>14</sup> con un número similar de pruebas de desplazamiento y corte (cuatro pruebas en total) en un simulador físico logró discriminar los mismos niveles de conocimiento y clasificar correctamente al 74% de los voluntarios participantes. La evaluación en ese estudio se realizó utilizando el OSATS. Usando el mismo simulador, Emper *et al.*<sup>15</sup>, pudo discriminar dos niveles básicos de competencia (novatos vs. Expertos) por medio de variables cinemáticas como longitud del recorrido, velocidad de retiro del instrumento, y «suavidad» de los movimientos. Fraser *et al.*<sup>16</sup>, también lograron discriminar dos niveles de habilidades psicomotoras (estudiantes y expertos) utilizando una colección de pruebas que incluían actividades de desplazamiento, habilidad con ambas manos, corte y realización nudos. Fraser *et al.* utilizaron un simulador físico básico mientras que Chmarra y Empel *et al.*, utilizaron un simulador físico con sistema de seguimiento al movimiento<sup>14</sup>.

El presente estudio también encontró que otras nueve pruebas de las 16 propuestas discriminan significativamente los grupos Novatos y Residentes, o los grupos Residentes y Expertos, y por tanto, pueden ser implementadas como complemento a las cuatro pruebas que efectivamente discriminan los tres grupos estudiados. Por ejemplo, dos de estas nueve pruebas evaluaron satisfactoriamente la actividad de desplazamiento de objetos, la cual no fue evaluada satisfactoriamente con las cuatro pruebas que discriminaron los tres grupos. Las nueve pruebas fueron

C4-»Malla», Di1-»Plástico», De2-»Cubos», De3-»Torre», C1-»Aluminio», C2-»Figura», C3-»Medio globo», S1-»Plastilina» y S4-»Collar».

La utilidad de esas nueve pruebas, sin embargo, estaría limitada a su empleo como parte de una colección de pruebas. Su empleo como pruebas independientes puede ser fuente de error en una evaluación de competencias. Por ejemplo, la prueba Di1-»Plástico» discriminó los grupos Residente y Experto pero el índice de desempeño fue más favorable, aunque no de manera significativa, para el grupo Residente que para el grupo Experto. Un problema similar fue visto con la prueba De3-»Torre».

La principal limitación de este estudio es el cálculo de algunos de los índices relativos de desempeño en las pruebas. Por ejemplo, la evaluación de la «precisión» en una tarea de corte fue realizada por medio de la medición de la máxima desviación del corte con respecto a una demarcación pre-establecida. No obstante, seleccionar el punto de máxima desviación y medirlo con una regla fue inadecuado y no adicionó objetividad a la prueba. Una redefinición de los índices puede hacerlos más sencillos de manejar y permite asegurar mayor objetividad en trabajos futuros.

## Conclusión

La colección propuesta de pruebas simples llevadas a cabo en un simulador físico permitió clasificar correctamente tres grupos de practicantes de CMI de acuerdo con su nivel de competencia: novato, residente y cirujano experto. Sin embargo, este resultado se confirma más en algunas de las pruebas propuestas que en otras. Es decir, solo cuatro de las pruebas permitieron clasificar correctamente a los practicantes en los tres grupos de estudio. No obstante, esas cuatro pruebas clasifican a los practicantes en solo dos de las habilidades básicas de CMI (disección y sutura). Esto

significa que persiste la necesidad de diseñar de pruebas simples en actividades de desplazamiento y corte que puedan complementar una evaluación psicomotora básica de los practicantes de CMI, y la redefinición de criterios de evaluación que permitan mejorar la objetividad de evaluación. Investigaciones futuras son necesarias para el desarrollo de evaluación y entrenamiento de cirujanos de CMI, particularmente a nivel regional.

#### Agradecimientos:

Los autores agradecen al Centro de Simulación Clínica de la Pontificia Universidad Javeriana y a su Director, Dr. Adalberto Amaya por permitirnos llevar a cabo este estudio en sus instalaciones

#### Conflicto de interés:

Los autores declaran no tener ningún conflicto de interés

#### Referencias

1. Cuschieri A. Laparoscopic surgery: Current status, issues and future developments. *Surg.* 2005; 3: 125-38.
2. Ruiz D, Pérez V, Betancur M, Bustamante J. Cirugía robótica mínimamente invasiva: análisis de fuerza y torque. *Rev Ing Biomédica.* 2010; 4:84-92.
3. Adrales GL, Chu UB, Witzke DB, Donnelly MB, Hoskins D, Mastrangelo MJ, *et al* . Evaluating minimally invasive surgery training using low-cost mechanical simulations. *Surg Endosc.* 2003; 17: 580-5.
4. Ritchie WP. Basic certification in surgery by the American Board of Surgery (ABS): what does it mean? Does it have value? Is it relevant? A personal opinion. *Ann Surg.* 2004; 239: 133-9.
5. Roberts KE, Bell RL, Duffy AJ. Evolution of surgical skills training. *World J Gastroenterol.* 2006; 12: 3219-24.
6. Naeem N. Validity, reliability, feasibility, acceptability and educational impact of direct observation of procedural skills (DOPS) *J Coll Physicians Surg Pak.* 2013; 23: 77-82.
7. Ghaderi I, Manji F, Park YS, Juul D, Ott M, Harris I, *et al* . Technical skills assessment toolbox: A review using the unitary framework of validity. *Ann Surg.* 2014; 261(2): 251-62.
8. Anaya-Prado R, Ortega LLH, Ramírez SME, Vázquez GJA, Medina PJB, Campos CSF, *et al* . Evaluación Objetiva de Habilidades Quirúrgicas. Modelo Mexicano. *Rev Latinoam Cir.* 2011; 1: 27-33.
9. Martin JA, Regehr G, Reznick R, Macrae H, Murnaghan J, Hutchison C, *et al* . Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg.* 1997; 84: 273-8.
10. Ahmed K, Miskovic D, Darzi A, Athanasiou T, Hanna GB. Observational tools for assessment of procedural skills: a systematic review. *Am J Surg.* 2011; 202: 469-80.
11. Halvorsen FH, Elle OJ, Fosse E. Simulators in surgery. *Minim Invasive Ther Allied Technol.* 2005; 14: 214-23.
12. Martinez AM, Espinoza DL. Novel laparoscopic home trainer. *Surg Laparosc Endosc Percutan Tech.* 2007; 17: 300-2.
13. Aggarwal R, Moorthy K, Darzi A. Laparoscopic skills training and assessment. *Br J Surg.* 2004; 91: 1549-58.
14. Chmarra MK, Klein S, de Winter JCF, Jansen F-W, Dankelman J. Objective classification of residents based on their psychomotor laparoscopic skills. *Surg Endosc.* 2010; 24: 1031-9.
15. Empel P, Rijssen L, Commandeur J, Verdam ME, Huirne J, Scheele F, *et al* . Validation of a new box trainer-related tracking device: the TrEndo. *Surg Endosc.* 2012; 26: 2346-52.
16. Fraser SA, Klassen DR, Feldman LS, Ghitulescu GA, Stanbridge D, Fried GM. Evaluating laparoscopic skills: setting the pass/fail score for the MISTELS system. *Surg Endosc.* 2003; 17: 964-7.