

ENTRENAMIENTO EN CIRUGÍA ROBÓTICA: VALIDACIÓN DE UNA ESCALA PARA LA EVALUACIÓN OBJETIVA DE DESTREZAS

RENATA SÁNCHEZ (1)
OMAIRA RODRÍGUEZ (2)
VALENTINA BÁEZ (3)
ORIANA SALAMO (3)
MARISA DI NATALE (3)
ALEXIS SÁNCHEZ ISMAYEL (4)

RESUMEN

Objetivo: El objetivo de este estudio consiste en determinar la capacidad de la escala GEARS (global evaluative assessment of robotic skills) para diferenciar individuos con distinto nivel de experiencia en cirugía robótica, como un parámetro fundamental para la validación de la misma.

Método: Es un estudio transversal que incluye a tres grupos de individuos con diferente nivel de experiencia en cirugía robótica (expertos, intermedios, novatos), a los cuales se les asignó una tarea sobre un modelo de entrenamiento, siendo evaluados mediante la escala en estudio (GEARS), aplicada por dos evaluadores. La diferencia entre los grupos se determinó mediante el método no paramétrico de Mann Whitney y la consistencia inter-observador se estudió mediante el coeficiente W de Kendall.

Resultados: Se realizaron un total de 15 evaluaciones, correspondientes a 5 individuos de cada grupo. La concordancia inter-observador de la escala GEARS fue de 0,96. El puntaje obtenido al realizar la evaluación fue de $29,8 \pm 0,4$; $24 \pm 2,8$ y 16 ± 3 ; para los expertos, intermedios y novatos respectivamente, siendo la diferencia entre ellos estadísticamente significativa. Todos los parámetros de la escala permiten diferenciar entre los distintos niveles de experiencia, con la excepción del ítem percepción de profundidad.

Conclusión: Se concluye que la escala GEARS demostró ser capaz de diferenciar entre individuos con distinto nivel de experiencia en cirugía robótica, validándose como una herramienta útil en el entrenamiento y evaluación del cirujano en formación.

Palabras clave: cirugía robótica, entrenamiento, GEARS.

TRAINING IN ROBOTIC SURGERY: VALIDATION OF A SCALE FOR OBJECTIVE SKILLS EVALUATION

ABSTRACT

Objective: The objective of this study is to determine the ability of the GEARS scale (global evaluative assessment of robotic skills) to differentiate individuals with different levels of experience in robotic surgery, as a fundamental validation of the same parameter. Method: This is a cross-sectional study that included three groups of individuals with different levels of experience in robotic surgery (expert, intermediate, novice) to which they were assigned a task on a training model, and was assessed by the GEARS scale applied by two reviewers. The difference between groups was determined by a nonparametric method from Mann Whitney and the consistency between the reviewers was studied by Kendall W coefficient.

Results: A total of 15 evaluations were performed to all the individuals in each group. The agreement between the reviewers of the scale GEARS was 0.96. The score of the assessment was 29.8 ± 0.4 to experts, 24 ± 2.8 to intermediates and 16 ± 3 to novices, with a statistically significant difference between all of them ($p < 0.05$). All parameters from the scale allow discriminating between different levels of experience, with exception of the depth perception item.

Conclusion: We conclude that the scale GEARS was able to differentiate between individuals with different levels of experience in robotic surgery. Therefore is a validated and useful tool to evaluate surgeons in training.

Keywords: robotic surgery, training, GEARS.

- 1 Profesor Agregado. Cátedra de Clínica y Terapéutica Quirúrgica "C", servicio de Cirugía III. Programa de Cirugía Robótica. Hospital Universitario de Caracas. MSVC
- 2 Profesora Asistente. Cátedra de Clínica y Terapéutica Quirúrgica "C", servicio de Cirugía III. Co-coordinador del Programa de Cirugía Robótica Hospital Universitario de Caracas. MSVC
- 3 Médico Cirujano. Programa de Cirugía Robótica. Hospital Universitario de Caracas.
- 4 Profesor Asociado. Cátedra de Clínica y Terapéutica Quirúrgica "C", servicio de Cirugía III. Magister Scientiarum en Ciencias Médicas. Coordinador del Programa de Cirugía Robótica. Hospital Universitario de Caracas. MSVC. SAGES.

La incorporación de nuevas tecnologías en la cirugía moderna ha traído consigo la necesidad de implementar cambios en la enseñanza de la práctica quirúrgica actual, siendo necesario dirigir el entrenamiento hacia ambientes seguros para poder garantizarles a nuestros pacientes una atención de calidad, con altas tasas de efectividad y baja morbilidad. La introducción de la cirugía laparoscópica proporciona grandes ventajas sobre la cirugía abierta^{1,3}; sin embargo, también es cierto que enfrenta al cirujano a dificultades propias de la técnica como lo son: pérdida de la percepción de profundidad producto de la visión en dos dimensiones, disminución en el rango de movimientos de los instrumentos, pérdida de la sensación táctil y el efecto fulcrum^{4,5}.

Entre los más recientes adelantos en el campo de la cirugía mínimamente invasiva se encuentra la incorporación de la tecnología robótica; el sistema Da Vinci® (Intuitive Surgical Inc.) es el único sistema disponible para la realización de cirugía laparoscópica asistida con robot, el cual dispone de una óptica binocular que aporta una óptima visión tridimensional y sus instrumentos diseñados especialmente con un sistema de múltiples poleas permiten gran maniobrabilidad con siete grados de libertad, lo que a su vez permite superar las dificultades propias del abordaje laparoscópico^{6,7}.

La enseñanza de la cirugía robótica debe iniciarse con una introducción teórica relacionada con los principios básicos, funcionamiento y estudio de los componentes del sistema, para luego pasar a la realización de prácticas dirigidas fundamentalmente al dominio de los instrumentos desde la consola⁸. Las prácticas deben ser distribuidas (con un patrón cronológico de entrenamiento), estructuradas (definiendo los objetivos y planificando los protocolos dirigidos para alcanzar las metas específicas)⁹ y deliberadas (relacionada con la formación de expertos en determinada área)^{9,10}.

Para alcanzar prácticas estructuradas y deliberadas es necesaria la incorporación de objetivos métodos de evaluación, que permitan definir el grado de destreza que se pretende alcanzar antes de pasar al siguiente nivel de instrucción, y que, adicionalmente, proporcionen al individuo en entrenamiento una adecuada retroalimentación, que se traduce en una mayor motivación por alcanzar las metas establecidas.

Entre los métodos de evaluación, el tiempo en el cual se realiza una tarea es uno de los parámetros más utilizados; sin embargo, múltiples estudios demuestran que la realización de determinada tarea en un tiempo menor no es reflejo de un adecuado dominio de la técnica, ya que podría realizarse con movimientos bruscos, numerosos y poco precisos¹¹, por lo que se hace necesaria la incorporación de escalas de evaluación que introduzcan elementos directamente relacionados con el dominio de la técnica.

Vassiliou y col. en el año 2005 propusieron la incorporación de una escala global (GOALS: global assessment of lapa-

roscopic skills) para la evaluación del desempeño en cirugía laparoscópica¹², la cual ha sido utilizada en múltiples estudios, demostrándose su confiabilidad, consistencia y validez, siendo capaz de diferenciar entre individuos con distinto nivel de experiencia en cirugía mínimamente invasiva¹³. Recientemente Goh y colaboradores del departamento de Urología del Baylor College of Medicine, en Houston-Texas, han propuesto una modificación de la escala original de GOALS, adaptando a ésta elementos propios de la cirugía robótica, para lo cual se incluyen los parámetros control robótico (relacionado con el dominio de los instrumentos y la cámara desde la consola) y control de la fuerza (ya que la falta de háptica puede conducir a ruptura de suturas o tejidos, siendo un factor directamente relacionado con morbilidad durante el cumplimiento de la curva de aprendizaje). Dicha escala creada según las modificaciones descrita ha sido denominada GEARs (global evaluative assessment of robotic skills), donde se evalúan seis parámetros, para un puntaje mínimo de 6 y un máximo de 3014 (Figura N° 1).

La incorporación de una escala como método de evaluación o certificación requiere de su validación previa. Uno de los elementos de validación más relevantes es la capacidad del método de evaluación para diferenciar entre sujetos con distintos niveles de destreza, es decir, de distinguir a aquellos individuos expertos de los novatos, lo que se denomina validez de construcción, y es el tipo de validación que se determinará en la presente investigación. Ésta constituye una de las características más valiosas de método de evaluación, ya que si no es posible detectar la diferencia entre los novatos y los expertos de forma inicial, las mejoras que los novatos realicen a lo largo del entrenamiento resultarían indetectables y no se podrían trazar metas en relación con el desempeño deseado¹⁵.

MÉTODOS

Se trata de un estudio de corte transversal, donde se tomaron 15 individuos, divididos en tres grupos basados en el nivel de experiencia en cirugía robótica: expertos, intermedios y novatos. En el grupo de expertos se incluyeron a individuos con experiencia superior a quince casos de cirugía robótica, el grupo intermedio se conformó por cirujanos que han recibido entrenamiento formal en el dominio de la consola del sistema Da Vinci® SHD, pero no han tenido participación en cirugías in vivo, y por último el grupo de novatos incluyó a individuos que no han tenido experiencia alguna en cirugía robótica. Muestras similares han sido utilizadas ampliamente en estudios de validación descritos en la literatura mundial y han demostrado ser suficientes para una adecuada validación de modelos, simuladores y escalas^{16,17}.

A todos los individuos se les explicó mediante una demostración la tarea que debían realizar. El ejercicio evaluado consiste en la realización de una sutura continua, con un punto y anuda-

Figura N° 1. ESCALA GEARS (global evaluative assessment of robotic skills) Goh et al. J Urol 2012 187:247-52

Percepción de Profundidad				
1	2	3	4	5
Constantemente sobrepasa el objetivo, movimientos amplios, corrige lentamente		Algunas fallas en la toma del objetivo, pero corrige		Dirige los instrumentos en el plano correcto hacia el objetivo
Destreza bimanual				
1	2	3	4	5
Usa solo una mano, ignora la mano no dominante, pobre coordinación entre ambas		Usa ambas manos, pero la interacción entre ambas no es óptima.		Usa ambas manos de manera complementaria para una óptima exposición
Eficiencia				
1	2	3	4	5
Muchos movimientos tentativos, cambios frecuentes en el paso a realizar, no progresa		Movimientos lentos, pero organizados y razonables		Confiado, eficiente, se mantiene enfocado en el objetivo.
Control de la fuerza				
1	2	3	4	5
Movimientos bruscos, desgarrar el tejido, daño a las estructuras. Ruptura frecuente de la sutura		Manejo razonable de los tejidos, ocurre daño menor. Ruptura ocasional de la sutura		Manejo adecuado de los tejidos, tracción apropiada de los mismos. Sin ruptura de la sutura
Autonomía				
1	2	3	4	5
Incapaz de terminar el procedimiento		Es capaz de terminar la tarea de maneja segura, con algo de guía por tutor.		Capaz de completar la tarea por sí solo, sin guía.
Control del robot				
1	2	3	4	5
No optimiza la posición de las manos en la consola, frecuente colisión. La visión no es óptima		Colisión ocasional de las manos. La visión es en ocasiones no óptima.		Control adecuado de la cámara. Posición óptima de las manos, sin colisión.

do inicial, seguido de tres pasadas sobre la incisión antes de realizar el nudo final (Figura N°2). El modelo utilizado se trata de una lámina de goma EVA (etileno vinyl acetato) colocada en forma tubular sobre la cual se ha realizado una incisión previamente. Este tipo de simulador inanimado ya ha sido previamente utilizado por nosotros en la evaluación de curvas de aprendizaje en cirugía robótica y corresponde con la tarea número nueve del programa de entrenamiento propuesto por Dulan y colaboradores, quienes han demostrado que se trata de un ejercicio con una elevada capacidad para diferenciar entre diferentes niveles de destreza en el uso del sistema robótico¹⁸.

Las tareas se llevaron a cabo en el quirófano de cirugía robó-

tica del Hospital Universitario de Caracas (Quirófano I). En el caso de los novatos, estos recibieron un corto curso introductorio en relación con los controles de la consola del Sistema Da Vinci® SHD.

Los ejercicios fueron grabados en formato digital para su posterior calificación por dos evaluadores, quienes no estaban en conocimiento del nivel de experiencia del individuo que estaba siendo evaluado. Los datos fueron almacenados en tablas diseñadas para tal fin en Excel (Microsoft Office®) y posteriormente analizados estadísticamente. El objetivo de usar dos evaluadores fue determinar la variabilidad interobservador, la concordancia entre estos fue evaluada utilizando el coeficiente W de Kendall.

La diferencia en el puntaje obtenido por los individuos de cada grupo (expertos, intermedios y novatos) se determinó usando el método no paramétrico de Mann Whitney con un nivel de significación de 5%. Se estudió tanto el puntaje global como cada ítem en particular, con la finalidad de determinar cuáles son los parámetros verdaderamente diferenciadores y la concordancia interna de la prueba, es decir, la similitud entre el resultado global y cada variable de la escala.

RESULTADOS

Se realizaron un total de 15 evaluaciones, correspondientes a cinco individuos de cada grupo (expertos, intermedios y novatos) por parte de cada uno de los evaluadores. La distribución por sexo y edad, así como el tiempo empleado y el puntaje obtenido por cada uno de los grupos se muestran en la tabla N° 1. La concordancia interobservador obtenida fue alta ($r=0,96$).

El tiempo empleado en la realización de la tarea fue de 311 ± 58 segundos para los expertos, 578 ± 198 " para los intermedios y 873 ± 282 " en el caso de los novatos. La diferencia fue estadísticamente significativa entre expertos e intermedios, ($p=0,02$), sin embargo, no fue así entre estos últimos y los novatos ($p=0,09$) (Figura N°3).

El puntaje promedio (GEARS) obtenido por cada grupo fue de $29,8 \pm 0,4$; $24 \pm 2,8$ y 16 ± 3 ; para los expertos, intermedios y novatos respectivamente. La diferencia entre cada uno de los grupos evaluados fue estadísticamente significativa, siendo el desempeño de los expertos superior a los intermedios ($p=0,008$), y éstos a su vez superiores a los novatos ($p=0,016$), con un nivel de significación de 5%. (Figura N°4).

El análisis individual de los parámetros de la escala evidenció que el ítem percepción de la profundidad no constituye un elemento diferenciador entre los individuos para ninguno de los tres niveles de experiencia, ya que todos obtuvieron la máxima cali-

Figura N°2. METODOLOGÍA

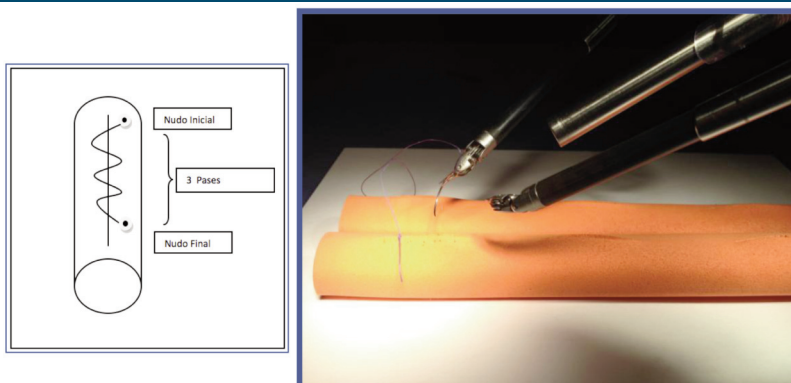


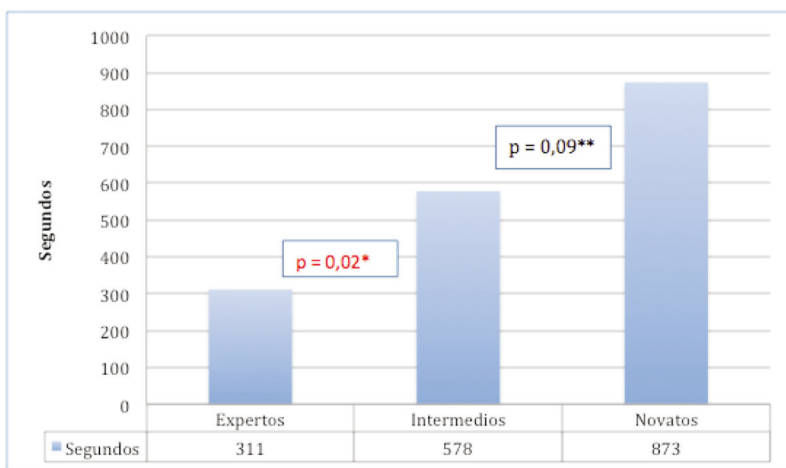
Tabla N° 1. Características de la muestra y evaluación según tiempo empleado en la tarea y escala GEARS.

	n	Sexo (H:M)	Edad (años)	Tiempo (segundos)	Puntaje GEARS*	
					Evaluador 1	Evaluador 2
Expertos	5	1:4	33,6 ± 4	311 ± 58	29,8 ± 0,4	29,8 ± 0,4
Intermedios	5	2:3	33,6 ± 6	578 ± 198	24 ± 2,8	23,6 ± 3,3
Novatos	5	4:1	25,8 ± 1	873 ± 282	16 ± 3	15,8 ± 2,9

Tabla de datos

* Concordancia interobservador (coeficiente W de Kendall) = 0,96

Figura N°3. Tiempo empleado para la realización de la tarea por cada uno de los grupos evaluados.



* Estadísticamente significativo ** No significativo

cación (5 puntos). La evaluación de la autonomía no permitió establecer una diferencia entre el desempeño de los expertos vs intermedios ($p=1,00$), mientras que todos los demás parámetros de la escala demostraron ser variables capaces de diferenciar entre individuos con diferente nivel de experiencia. (Tabla N°2).

DISCUSIÓN

La incorporación de la tecnología robótica a la cirugía mínimamente invasiva ha sido un gran avance al permitir superar limitaciones propias de la técnica, lo cual ha llevado a considerarla como una excelente opción para numerosos procedimientos, pero a su vez requiere de un adecuado entrenamiento del equipo quirúrgico como factor fundamental para asegurar procedimientos efectivos y con baja mortalidad^{8,19}. El aprendizaje de la cirugía mínimamente invasiva debe ser programado y llevado a cabo en laboratorios diseñados para tal fin, valiéndose de la simulación⁹.

Las prácticas supervisadas en ambientes reales o las cirugías tutoradas ya no son consideradas como la primera opción, ya que consumen tiempo, aumentan los costos y tienen obvias implicaciones médico legales. Los cirujanos en formación deben adquirir las habilidades en laboratorios diseñados para tal fin, que permitan al docente enfocarse en puntos clave sin poner en riesgo el bienestar del paciente. Múltiples estudios demuestran que las habilidades aprendidas fuera del quirófano son efectivamente transferidas a las cirugías in vivo, permitiendo avanzar en la curva de aprendizaje²⁰. La práctica dirigida a la formación de especialistas altamente capacitados debe ser distribuida y estructurada; haciendo especial énfasis en la realización de prácticas deliberadas en las cuales se realizan sesiones supervisadas donde se identifican necesidades específicas de cada individuo requeridas para el logro de las metas establecidas, teniendo como objetivo la formación de expertos. La práctica deliberada implica no solo la adquisición de una destreza, sino el perfeccionamiento de la misma^{9,10}.

La evaluación con parámetros objetivos permite realizar una retroalimentación inmediata, lo cual es considerado como un importante elemento en el desarrollo de habilidades,

y definitivamente su aplicación se traduce en una mayor motivación para los cirujanos en entrenamiento, a la vez que permite determinar el progreso en la adquisición de habilidades en la medida que se realizan las prácticas sucesivas, permitiendo establecer una meta en el nivel de destrezas.

El grupo de expertos obtuvo una puntuación mayor que los

intermedios y novatos, observándose además una mínima variabilidad entre los individuos del grupo (29,8±0,4). Esta homogeneidad de los datos es característica de los grupos entrenados, y difiere claramente de los grupos con menor experiencia. El análisis de los datos registrados comprueba que la puntuación obtenida mediante la escala GEARS es capaz de diferenciar distintos niveles de experiencia en cirugía robótica, lo cual la convierte en una herramienta útil para evaluación y retroalimentación durante el cumplimiento del programa de entrenamiento, a la vez que permite monitorear la progresión en la adquisición de destrezas y determinar metas a alcanzar para cada nivel de entrenamiento.

Los resultados indican que la escala tiene una alta confiabilidad, con una concordancia interobservador de r=0.96. De manera general, se acepta que un valor superior a r=0,8 cataloga al método como una herramienta objetiva no dependiente del evaluador¹⁵.

La confiabilidad interna con la mayoría de los parámetros que la constituyen fue excelente, sin embargo, la variable percepción de profundidad demostró no ser un elemento capaz de diferenciar entre individuos con diferente nivel de experiencia, lo cual se explica por el hecho de que la óptima visión tridimensional aportado por el sistema binocular del sistema Da Vinci® permite, aún a los cirujanos poco entrenados, tener una excelente puntuación en este ítem. Así mismo, en el caso del parámetro automatismo, es decir, la capacidad para terminar la tarea por sí solo, no demostró diferencia entre los niveles intermedio y expertos, probablemente debido a que la gran versatilidad del sistema robótico permite a individuos medianamente entrenados completar la tarea con una mínima instrucción.

Debido a lo descrito, pareciera sensato omitir el parámetro percepción de profundidad de la escala, convirtiéndola en una escala de cinco ítems con una puntuación máxima de 25 puntos, la cual pudiéramos denominar GEARS modificada. La evaluación de esta modificación de la escala propuesta formará parte de futuros protocolos de esta línea de investigación dentro del marco del programa de cirugía robótica del Hospital Universitario de Caracas y la Facultad de Medicina de la Universidad Central de Venezuela.

La escala GEARS demostró ser capaz de diferenciar entre

Figura N°4. Evaluación del desempeño mediante la escala GEARS

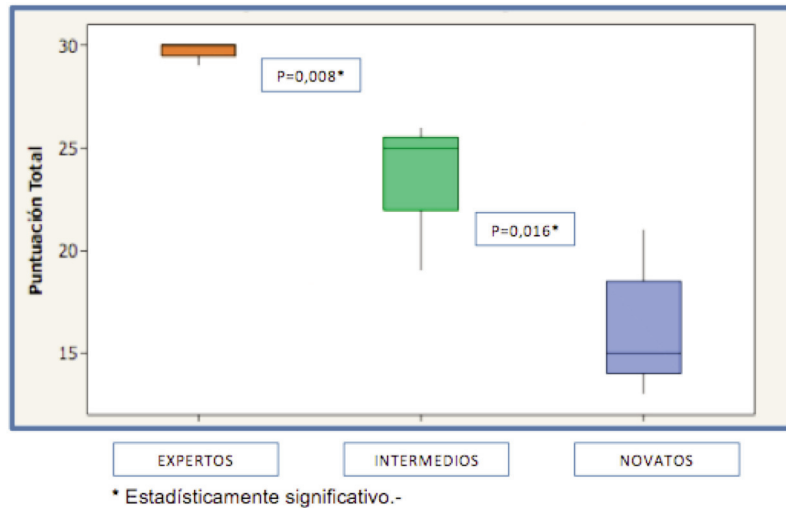


Tabla N° 2. Análisis individual de la capacidad de cada parámetro de la escala GEARS para diferenciar entre individuos de primer nivel de experiencia

Parámetro	E	=	I	=	N
Percepción de la profundidad	E	=	I	=	N
Destreza bimanual	E	>	I	>	N
Eficiencia	E	>	I	>	N
Control de la fuerza	E	>	I	>	N
Autonomía	E	=	I	>	N
Control del robot	E	>	I	>	N

E: EXPERTOS, I: INTERMEDIOS, N: NOVATOS

E : Parámetros no diferenciadores entre los grupos.-

individuos con distinto nivel de experiencia en cirugía robótica, validándose como una herramienta útil en el entrenamiento y evaluación del cirujano en entrenamiento.

REFERENCIAS

1. Perissat J, Collet D, Belliard R, Desplantez J, Magne E. Laparoscopic cholecystectomy. The state of the art. A report on 700 consecutive cases. World J Surg 1992; 16(6): 1074-1082.
2. Guller U, Hervey S, Purves H, Muhlbauer L, Peterson E, Eubank S, et al. Laparoscopic versus open appendectomy: outcomes comparison based on a large administrative database. Ann Surg 2004; 239(1):43-52.
3. Staudacher C, Vignali A. Laparoscopic surgery for rectal cancer: The state of the art. World J Gastrointest Surg 2010; 2(9):275-282.

4. Fraser, Klassen R, Feldman D, Ghitulescu D, Stanbridge D, Fried G. Evaluating laparoscopic skills. *Surg Endosc* 2003; 17(6):964-967.
5. Aggarwal R, Moorthy K, Darzi A. Laparoscopic skills training and assessment. *Br J Surg* 2004; 91(12):1549-1558.
6. Corcione F, Esposito C, Cucurullo D, Settembre A, Miranda N, Amato F, et al. Advantages and limits of robot-assisted laparoscopic surgery. *Surg Endosc* 2005; 19(1):117-119.
7. Jayaraman S, Quan D, Al-Ghamdi I, El-Deen F, Schlachta C. Does robotic assistance improve efficiency in performing complex minimally invasive surgical procedures? *Surg Endosc* 2010; 24(3):584-588.
8. Herrom D, Marohn M; SAGES-MIRA Robotic surgery consensus group. A consensus document on robotic surgery. *Surg Endosc* 2008; 22(2):313-325.
9. Tsuda S, Scott D, Doyle J, Jones D. Current Problems in Surgery: Surgical Skills Training and Simulation. *Curr Probl Surg* 2009; 46(4): 271-370.
10. Ericsson K, Krampe R, Tesch-Romer. The role of deliberated practice in the acquisition of expert performance. *Psychol Rev* 1993;100(3):363-406.
11. Mason J, Ansell J, Warren N, Torkington J. Is motion analysis a valid tool for assessing laparoscopic skill? *Surg Endosc* 2013; 27(5):1468-1477.
12. Vassiliou M, Feldman L, Andrew C, Bergman S, Leffondre K, Stanbridge D, et al. A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg* 2005; 190(1):107-113.
13. Chang L, Hogle N, Moore B, Graham M, Sinanan M, Bailey R, et al. Reliable assessment of laparoscopic performance in the operating room using videotaped analysis. *Surg Innov* 2007; 14(2):122-126.
14. Goh A, Goldfarb D, Sander J, Miles B, Dunkin B. Global evaluative assessment of robotic skills: a clinical assessment tool to measure robotic surgical skills. *J Urol* 2012; 187:247-252.
15. McDougall E. Validation of Surgical Simulators. *J Endourology* 2007; 21(3):244-247.
16. Van Empel P, Rijssen L, Commandeur J, Verdam M, Huirne J, Scheele F, et al. Validation of a new box trainer-related tracking device: the TrEndo. *Surg Endosc* 2012; 26(8):2346-2352.
17. Sánchez A, Otaño N, Rodríguez O, Sánchez R, Benítez G, Schweitzer M. Laparoscopic common bile duct exploration four-task training model: construct validity. *JSL* 2012;16(1):10-15.
18. Dulan G, Rege R, Hogg D, Gilberg-Fisher K, Arain N, Tesfay S, et al. Proficiency-based training for robotic surgery: construct validity, workload and expert levels for nine inanimate exercises. *Surg Endosc* 2012;26(6): 1516-1521.
19. Sachdeva A, Russell T. Safe Introduction of new Procedures and Emerging technologies in Surgery: Education, Credentialing and Privileging. *Surg Clin N Am* 2007;87:853-866.
20. Sturm L, Windsor J, Cosman P, Cregan P, Hewett P, Maddern G. A systematic review of skills transfer after surgical simulation training. *Ann Surg* 2008; 248(2):166-179.