

Evaluación objetiva de las habilidades técnicas en cirugía

Francisco Navarro Subiabre¹, Mauricio Gabrielli Nervi², Julian Varas Cohen²

Resumen

Introducción: La enseñanza tradicional de la cirugía se ha basado en la observación y retroalimentación directa por parte de un experto. Esto corresponde a una evaluación subjetiva y requiere de un tiempo mínimo de exposición no siempre fácil de obtener. Los resultados adversos y complicaciones se traducen en gastos innecesarios para los servicios de salud. Surge la necesidad de nuevos métodos de evaluación objetiva de las habilidades quirúrgicas. Se presenta una revisión de la literatura respecto a los métodos objetivos de evaluación de las habilidades técnicas en cirugía que actualmente existen y dar a conocer una nueva metodología de evaluación actualmente en desarrollo. **Métodos:** Se realizó una revisión de la literatura, utilizando PubMed, Google Scholar y Scielo. Se incluyeron estudios en inglés y español que utilizaron métodos objetivos estandarizados de evaluación de técnicas quirúrgicas, ya sea con técnica abierta o laparoscópica, además de revisiones de la literatura. **Resultados:** Se incluyeron las escalas OSATS, GOALS, GAGES, métodos basados en simulación y otros aún en desarrollo. Las escalas de evaluación han sido validadas en un gran número de procedimientos. La simulación en el área quirúrgica ha permitido generar una evaluación y retroalimentación previa al escenario operatorio real, reduciendo la tasa de errores e incrementando la eficiencia en éste. La Caja Negra es un método novedoso que incluye una gran cantidad de variables, permitiendo detectar pequeñas variaciones o errores en la ejecución de la técnica, lo que permite una evaluación más certera. **Conclusiones:** Múltiples métodos de evaluación objetiva de las habilidades técnicas han sido desarrollados y validados. Se espera que exista una incorporación sistemática de estas herramientas en los programas de formación de cirujanos de nuestro país.

Palabras clave: Cirugía; Habilidades; Evaluación

Abstract

Introduction: The traditional teaching of surgery has been based on observation and direct feedback from an expert. This is a subjective evaluation and requires a minimum exposure time not always easy to obtain. Adverse outcomes and complications translate into unnecessary expenses for health services. The need for new methods of objective evaluation of surgical skills arises. We present a review of the literature regarding the objective methods of evaluation of the technical skills in surgery that currently exist and present a new evaluation methodology currently under development. **Methods:** A review of the literature was performed, using PubMed, Google Scholar and Scielo. Studies in English and Spanish were included that used objective standardized methods of evaluation of surgical techniques, either with open or laparoscopic technique, in addition to reviews of the literature. **Results:** The scales OSATS, GOALS, GAGES, simulation-based methods and others still in development were included. The evaluation scales have been validated in a large number of procedures. The simulation in the surgical area has allowed to generate an evaluation and previous feedback to the real operating scenario, reducing the error rate and increasing the efficiency in it. The Black Box is a novel method that includes a large number of variables, allowing to detect small variations or errors in the execution of the technique, which allows a more accurate evaluation. **Conclusions:** Multiple methods of objective evaluation of technical skills have been developed and validated. It is expected that there is a systematic incorporation of these tools in the training programs of surgeons in our country.

Keywords: Surgery; Skills; Assessment

Fecha de envío: 18 de febrero de 2018 - Fecha de aceptación: 14 de septiembre de 2018

Introducción

Por muchos años la enseñanza de la cirugía ha girado en torno al modelo profesor-alumno, donde mediante la tutoría personalizada

el residente adquiere las competencias necesarias para su desempeño, siendo evaluado por un observador no siempre entrenado para hacerlo (en inglés "see one, do one"). Uno de los ejes de este proceso

(1) Escuela de Medicina, Facultad de Medicina, Pontificia Universidad Católica de Chile.

(2) Departamento de Cirugía Digestiva, Facultad de Medicina, Pontificia Universidad Católica de Chile.

Autor de Correspondencia: mgabrie@med.puc.cl



se basa en la observación directa de los distintos procedimientos que permiten incorporar las habilidades y al mismo tiempo generar una adecuada retroalimentación hacia el estudiante (Healey *et al.*, 2006). Para esto, es requerido habitualmente la repetición sistemática de un número determinado de intervenciones que indiquen “tiempo” de exposición permitiendo incorporar definitivamente dichas habilidades. Lamentablemente, el limitado volumen de pacientes a los que se exponen los residentes durante su formación ha disminuido paulatinamente en los últimos años, aumentando la dificultad en la adquisición de las distintas técnicas quirúrgicas (Sachs & Pawlik, 2015). Además, evidentemente estos métodos se supeditan en una evaluación subjetiva de las habilidades del residente por parte del observador, lo que no permite generar un entrenamiento sistematizado y estructurado (Grantcharov & Reznick, 2008). Por otro lado, es claro que los altos costos asociados a eventos adversos y complicaciones en el perioperatorio se traducen en un significativo aumento en los gastos hospitalarios (Garbens *et al.*, 2017). Es por esto, que en los últimos años ha ido surgiendo paulatinamente la tendencia hacia un cambio en este paradigma educativo basado en el tiempo, hacia uno basado en las competencias, como lo han sugerido distintas agrupaciones a nivel internacional. Esto último ha sido difícil de definir, principalmente por el escaso consenso que existe en torno al concepto de “competencia técnica”, así como los métodos para su correcta evaluación. El interés por un método de evaluación objetiva de las habilidades quirúrgicas de los residentes y de los cirujanos ha crecido exponencialmente en los últimos años.

Numerosos métodos de evaluación objetiva de habilidades técnicas han sido propuestos, sin embargo, su aplicación directa en los programas de formación quirúrgica ha sido escasa (Aggarwal *et al.*, 2007; Grantcharov & Reznick, 2008). El objetivo del presente estudio es realizar una revisión de la literatura respecto a los métodos objetivos de evaluación de las habilidades técnicas en cirugía que existen y dar a conocer una nueva metodología de evaluación actualmente en desarrollo.

Métodos

Se realizó una revisión de la literatura, utilizando PubMed, Google Scholar y Scielo. Se utilizaron las siguientes palabras claves: surgical OR operative OR technical OR laparoscopic AND skills OR competence AND objective assesment (PubMed, Google Scholar), y evaluación, habilidades, cirugía (Scielo). Se incluyeron estudios en inglés y español que se basaron en métodos objetivos estandarizados de evaluación de técnicas quirúrgicas, ya sea con técnica abierta o laparoscópica, además de revisiones de la literatura.

Respecto a los métodos analizados es necesario definir ciertos conceptos. Validez es definida como la “propiedad de ser verdadero, correcto y en conformidad con la realidad” (Gallagher *et al.*, 2003), y se compone de varias dimensiones, donde destacan principalmente la validez de contenido, validez de constructo y validez discriminante. La validez de contenido apunta a si los componentes del instrumento reflejan las habilidades que fueron designadas a medir. La validez de constructo se refiere a si este instrumento realmente mide el rasgo o la característica que debería medir; por ejemplo, un instrumento tendrá este tipo de validez si los residentes de mayor rango obtienen puntajes más elevados que los novatos.

La confiabilidad se refiere a si una herramienta es consistente con su resultado. Dentro de los elementos que la componen está la consistencia interna y la confiabilidad inter-evaluador. Este último punto es mejor evaluado cuando los examinadores desconocen la identidad y el nivel de entrenamiento del individuo. Se representa habitualmente con un coeficiente que va de 0,0 a 1,0, siendo mayor a 0,8 considerado como aceptable (Beard *et al.*, 2011).

Finalmente, la evaluación puede ser formativa o sumativa. Una evaluación formativa es aquella que se realiza con el objetivo de realizar retroalimentación al aprendiz y a la vez monitorizar su progreso. La sumativa, en cambio, es usada para la selección de unos aprendices sobre otros, por ejemplo, para algún tipo de certificación y por lo tanto requiere de resultados o metas predefinidas a alcanzar. Habitualmente este tipo de evaluación requiere estándares de validez y confiabilidad más elevados (van Hove PD *et al.*, 2010).

Resultados

Escalas de Evaluación General

OSATS

A fines de los años 90, Martin y cols. (Martin *et al.*, 1997), desarrollaron en la Universidad de Toronto una escala denominada *Objective Structured Assessment of Technical Skills* (OSATS). Se basa en la observación directa de los residentes o cirujanos realizando distintos procedimientos quirúrgicos en modelos de simulación. Se utilizaron 8 estaciones de 15 minutos de duración para la evaluación. La escala original contiene una primera parte que incluye un *checklist* dicotómico de acciones específicas para cada procedimiento (Tabla 1) y luego una escala de evaluación global (EEG), común a todas las estaciones utilizando una escala de Likert de 1 a 5 (Tabla 2).

Tabla 1: Checklist usado en OSATS para control de hemorragia y reparación de vena cava inferior. Se muestran los distintos ítems a evaluar en ambos pasos del procedimiento. Adaptado de Martin et al. (1997).

ITEM	No hecho/ Hecho incorrectamente	Hecho correctamente
<u>Control de hemorragia</u>		
1. Aplica presión para parar sangrado primero	0	1
2. Pide asistencia para succión en campo	0	1
3. Inspecciona la lesión liberando cuidadosamente la vena cava inferior	0	1
4. Asegura que todo el equipamiento necesario esté a la mano antes de comenzar	0	1
5. Control del punto de sangrado	0	1
<u>Reparación</u>		
6. Selecciona la sutura apropiada (4,0/5,0/6,0 polipropileno)	0	1
7. Selecciona el portaobjetos adecuado (vascular)	0	1
8. Selecciona la pinza adecuada (de Bakey)	0	1
9. Aguja cargada a 1/2-2/3 de la punta el 90% del tiempo	0	1

Tabla 2: Escala de evaluación global usada en OSATS. Se muestran los elementos evaluados y las definiciones utilizadas en la escala de puntuación. Adaptado de Martin et al. (1997).

Respeto por el tejido	1 Frecuentemente usó fuerza innecesaria sobre el tejido o causó daño por el uso inapropiado de instrumentos	2	3 Manejo cuidadoso del tejido, pero ocasionalmente causó daño inadvertido	4	5 Manipuló tejidos en forma adecuada, con un daño mínimo
Tiempo y movimiento	1 Muchos movimientos innecesarios	2	3 Relación tiempo/movimiento eficiente, pero con algunos movimientos innecesarios	4	5 Economía de movimientos y eficiencia máxima
Manejo instrumental	1 Realizó movimientos tentativos o extraños con instrumentos	2	3 Uso de instrumentos competente, aunque en algunos momentos pareció estar incómodo o rígido	4	5 Movimientos fluidos con instrumentos, sin incomodidades
Conocimiento de los instrumentos	1 Frecuentemente solicitó el instrumento equivocado o utilizó un instrumento inapropiado	2	3 Conoce los nombres de la mayoría de los instrumentos y usó el instrumento apropiado para la tarea	4	5 Claramente familiar con los instrumentos requeridos y sus nombres
Uso de asistentes	1 Constantemente ubicó a sus asistentes en forma deficiente o falló en usarlos	2	3 Buen uso de sus asistentes la mayoría del tiempo	4	5 Utilizó estratégicamente a sus asistentes para aprovecharlos al máximo
Flujo de operación y planificación anticipada	1 Frecuentemente se detiene durante la operación o necesita discutir el siguiente movimiento	2	3 Demostró habilidad para planificación anticipada con una progresión constante del procedimiento	4	5 Claramente planificó el curso de la operación con fluidez de un movimiento al siguiente
Conocimiento de procedimiento específico	1 Conocimiento deficiente. Necesitó instrucciones específicas en la mayoría de los pasos	2	3 Sabía todos los aspectos importantes de la operación	4	5 Demostró estar familiarizado con todos los aspectos de la operación

La confiabilidad y validez de constructo ha sido demostrada tanto para el *checklist* como también para la EEG (Niitsu *et al.*, 2013). Esto viene dado principalmente por la obtención de mayores puntajes por parte de los residentes en cursos más avanzados respecto a los recién ingresados.

Respecto a las limitaciones de esta escala, en primer lugar, en la sección de *checklist* específico para cada procedimiento surge el problema de la variabilidad que existe en el propio desarrollo de cada uno de los ítems a evaluar. En las descripciones originales de la OSATS, estos *checklist* fueron revisados y evaluados por diez cirujanos distintos, no involucrados directamente en la evaluación de los procedimientos (Martin *et al.*, 1997). Sin embargo, en la gran mayoría de los estudios no se reporta un proceso definido para el desarrollo de esta sección, lo que puede interferir con la validez de los resultados reportados en esta sección (Hatala *et al.*, 2015). Al contrario, las EEG se han mantenido constantes en los estudios con respecto a la descripción original, desarrollada y revisada por un grupo de expertos, por lo que no existirían diferencias significativas al respecto que pudiesen interferir con la validez de esta sección.

Otro punto, es el método de selección de los evaluadores, el cual es escasamente descrito en los estudios de validez de esta escala, al punto de ser sólo mencionar el cargo académico, lo que podría influir también en la evaluación de los participantes (Hatala *et al.*, 2015).

Sin embargo, pese a estas limitaciones, actualmente OSATS se utiliza como método de evaluación para cirujanos y residentes en donde analistas expertos y certificados revisan la cirugía entregando un puntaje final del desempeño técnico del operador. Estos analistas expertos han demostrado una alta correlación, superior al 75%, al evaluar en forma ciega los mismos procedimientos, ya que actualmente existe un método de entrenamiento objetivo de analista quirúrgico en la Universidad de Toronto (dato no publicado).

GOALS

En las últimas décadas, la cirugía laparoscópica se ha transformado en la primera línea en múltiples procedimientos y junto con este crecimiento, surge paralelamente la necesidad de una evaluación objetiva de estas habilidades. La escala *Global Operative Assessment of Laparoscopic Skills* (GOALS) fue desarrollada por Vassiliou y cols. (Vassiliou *et al.*, 2005) en la Universidad de McGill en Montreal y evalúa la ejecución de un procedimiento en cirugía mínimamente invasiva en cinco ítems (percepción de profundidad, destreza bimanual, eficiencia, manejo de tejidos, y autonomía). Consiste en una escala de Likert de 1 a 5, incluyendo una descripción en los puntajes 1, 3 y 5 para cada ítem de la intervención laparoscópica (Tabla 3). Posteriormente, se obtiene el puntaje total que equivale a una objetivación del desempeño del operador en la ejecución del procedimiento.

Tabla 3: Tabla original de GOALS. Se muestran los 5 ítems evaluados en la intervención y pequeñas descripciones para orientar la escala de puntuación. Adaptada de Vassiliou *et al.* (2005).

Percepción de profundidad	
1.	Constantemente sobrepasa el objetivo, balanceo amplio, lento para corregir.
2.	
3.	Algunas veces sobrepasa o pierde el objetivo, pero corrige rápidamente.
4.	
5.	Dirige con precisión instrumentos en el plano correcto al objetivo.
Destreza bimanual	
1.	Utiliza sólo una mano, ignora la mano no dominante, escasa coordinación entre ambas manos.
2.	
3.	Utiliza ambas manos, pero no optimiza la interacción entre ellas.
4.	
5.	Utiliza las manos de forma complementaria para proporcionar una exposición óptima
Eficiencia	
1.	Esfuerzos inciertos e ineficientes; muchos movimientos tentativos; cambiando constantemente de foco o persistiendo sin progreso.
2.	
3.	Lento, pero los movimientos planificados están razonablemente organizados.
4.	
5.	Conducta segura y eficiente, mantiene el foco en la tarea hasta que ésta se realiza mejor mediante un enfoque alternativo.
Manejo de tejidos	
1.	Movimientos bruscos, desgarrar el tejido, daña las estructuras adyacentes, control deficiente de la pinza, con frecuencia se desliza el tejido sostenido.
2.	
3.	Maneja los tejidos razonablemente bien, traumatismo menor en el tejido adyacente (es decir, sangrado o deslizamiento innecesario ocasional de la pinza).
4.	
5.	Maneja bien los tejidos, aplica la tracción adecuada, lesión insignificante a estructuras adyacentes.
Autonomía	
1.	Incapaz de completar la tarea completa, incluso con orientación verbal.
2.	
3.	Capaz de completar la tarea en forma segura con orientación moderada.
4.	
5.	Capaz de completar la tarea en forma independiente, sin preguntar.

Los estudios iniciales comenzaron validando esta escala en procedimientos simples como la disección de la vesícula del hígado (Vassiliou *et al.*, 2005), para posteriormente incluir a la colecistectomía y la apendicectomía laparoscópica (Gumbs *et al.*, 2007). Recientemente, también fue validado para procedimientos más complejos como la reparación de hernia inguinal (Kurashima *et al.*, 2011) y hernia incisional (Vaillancourt *et al.*, 2011). Finalmente, Hogle y cols. (Hogle *et al.*, 2014) la aplicaron en una cohorte pequeña de residentes de subespecialidad para evaluar procedimientos de mayor complejidad como funduplicatura de Nissen, banda gástrica ajustable y bypass gástrico en Y de Roux, siendo capaces de demostrar la validez de constructo de esta escala al diferenciar adecuadamente a los novatos de los más experimentados residentes.

Pese a que otros métodos han surgido para intentar evaluar objetivamente las habilidades en procedimientos laparoscópicos, su aplicación generalizada ha sido entorpecida por la complejidad de las herramientas de evaluación. Uno de ellos es un sistema de puntuación que permitía diferenciar a residentes novatos de los más avanzados, sin embargo, requería de 3 cirujanos experimentados revisando videos de colecistectomías laparoscópicas y asignando puntajes en 23 pasos distintos de ésta (Eubanks *et al.*, 1999). Otra herramienta consistía en una revisión de videos que al igual que la anterior también demostró validez de constructo, pero el análisis de cada tarea era específico para colecistectomía laparoscópica, no siendo aplicable para otros procedimientos (Sarker *et al.*, 2006). GOALS puede ser utilizada para identificar las áreas donde el residente presente debilidades y requiera un tiempo mayor de práctica. De esta forma, es posible generar una retroalimentación adecuada entre éste y el evaluador que permita dirigir el entrenamiento hacia dichas áreas, lo que ha demostrado ser una herramienta útil para un mejor desempeño de los residentes (Grantcharov *et al.*, 2007). Además, mediante esta escala es posible evaluar el rendimiento de cada residente respecto al desempeño promedio de sus propios compañeros para cada procedimiento evaluado (Gumbs *et al.*, 2007).

GAGES

Con el desarrollo de nuevas tecnologías, las técnicas endoscópicas han ocupado un papel protagónico en los programas de formación de cirujanos generales y de especialidades. La simulación mediante realidad virtual ha formado parte importante de la enseñanza de estas habilidades, sin embargo, sigue existiendo un área no menor que involucra el aprendizaje basado en el paciente y el tiempo de exposición a estas técnicas habitualmente varía mucho entre los distintos programas. El método de evaluación de estas competencias no está claramente definido y el número de procedimientos ejecutados dista de ser un parámetro objetivo para determinar el nivel de competencia de los residentes en esta área. La escala Global Assessment of Gastrointestinal Endoscopic Skills (GAGES)

se desarrolló en base a las habilidades fundamentales tanto para endoscopia digestiva alta (GAGES-UE), como para colonoscopia (GAGES-C). Siguiendo la misma modalidad de escala de Likert que GOALS, ésta evalúa los ítems "intubación del esófago", "alcance de la navegación", "habilidad para mantener un campo endoscópico limpio",

Tabla 4: Escala GAGES-UE original, utilizada en endoscopia digestiva alta. Se muestran los 5 ítems evaluados en el procedimiento y sus respectivas descripciones para orientar la evaluación. Adaptada de Vassiliou *et al.* (2010).

Intubación del esófago	
5.	Capaz de intubar esófago en forma independiente, sin malestar en paciente.
4.	
3.	Requiere indicaciones y pautas detalladas.
2.	
1.	Incapaz de intubar correctamente requiriendo hacerse cargo.
Alcance de la navegación	
5.	Perfectamente capaz de manipular el alcance en el tracto digestivo alto en forma autónoma.
4.	
3.	Requiere orientación verbal para navegar completamente por el tracto digestivo alto.
2.	
1.	No es capaz de conseguir objetivos a pesar de detalladas indicaciones verbales, requiriendo hacerse cargo.
Habilidad para mantener un campo endoscópico limpio	
5.	Usa insuflación, succión e irrigación en forma óptima para mantener una vista clara del campo endoscópico.
4.	
3.	Requiere indicaciones moderadas para mantener una vista clara.
2.	
1.	Incapacidad para mantener una vista clara a pesar de las extensas indicaciones verbales.
Instrumentación (si es aplicable, dejar en blanco si no)	
5.	Dirige perfectamente el instrumento al objetivo deseado.
4.	
3.	Requiere alguna orientación y/o múltiples intentos para dirigir el instrumento al objetivo.
2.	
1.	Incapaz de dirigir instrumento al objetivo pese al entrenamiento.
Calidad del examen	
5.	Completa el examen eficiente y cómodamente.
4.	
3.	Requiere asistencia moderada para llevar a cabo un completo y cómodo examen.
2.	
1.	No puede realizar un examen satisfactorio a pesar de la asistencia verbal y manual requiriendo hacerse cargo del procedimiento.

“instrumentación” (por ejemplo, para evaluar la toma de biopsias o de dirigir un determinado instrumento hacia un objetivo) y “calidad del examen” (Tabla 4). Esta escala permite al alumno obtener una retroalimentación dirigida hacia los aspectos en los que presenta mayor debilidad a la hora de realizar el procedimiento. La validez de esta escala ha sido evaluada en algunos estudios recientes. Vassiliou y cols. (Vassiliou *et al.*, 2010), obtuvieron una confiabilidad de 0,89 y 0,95 para GAGES-UE y GAGES-C, respectivamente.

Simulación

En los últimos años, el campo de la simulación ha cobrado cada vez una mayor importancia en el desarrollo y en múltiples áreas del conocimiento y la industria, con el objetivo de maximizar las condiciones de seguridad, minimizando los riesgos asociados. Dentro del campo de la educación médica, las demandas por una formación técnica orientada a la seguridad del paciente, obteniendo lecciones aprendidas de otras industrias como la aeronáutica y militar, han puesto sobre la mesa nuevas responsabilidades para los profesionales de la salud en formación. Respondiendo a estas inquietudes, en las últimas tres décadas, la simulación se ha convertido en una herramienta fundamental en el entrenamiento de profesionales de la salud, en la optimización en la calidad del cuidado y en los estándares de seguridad hacia paciente (Moya *et al.*, 2017).

Por otro lado, la simulación se ha hecho también cada vez más presente en los programas de formación de cirujanos, desarrollando habilidades que han permitido reducir las curvas de aprendizaje para nuevos procedimientos, incrementar la eficiencia en el escenario operatorio y disminuir la tasa de errores (Aggarwal *et al.*, 2007). En este sentido, uno de los ejes para un entrenamiento simulado satisfactorio corresponde a la presencia de una adecuada evaluación y retroalimentación por parte de un tutor experto (Shippey, 2011).

La tendencia de los resultados apunta a una mejora en las habilidades técnicas de los participantes sometidos a entrenamiento basado en simulación respecto a los grupos no entrenados, traducándose principalmente en menor tiempo operatorio, menos errores durante el procedimiento y menor incomodidad para el paciente (Buckley *et al.*, 2014). Se ha visto que esta mejora en las habilidades técnicas puede mantenerse incluso por hasta 5 meses en ausencia de práctica (Stefanidis & Acker, 2008), y que ésta no se correlaciona con la dificultad del entrenamiento (Stefanidis *et al.*, 2007).

La transferencia de habilidades desde el escenario simulado hacia el escenario operativo real (EOR) ha sido también tratada en la literatura. Korndorferr y cols (Korndorferr *et al.*, 2005), aleatorizaron a un grupo de residentes en dos subgrupos para realizar una funduplicatura de Nissen mediante sutura laparoscópica. El primer

grupo fue entrenado en un modelo de sutura mediante video y el segundo no recibió entrenamiento alguno. Posteriormente, ambos grupos llevaron a cabo el procedimiento en un modelo porcino. Ambos demostraron un desempeño significativamente mayor al basal evaluado al inicio, sin embargo, el grupo entrenado obtuvo un mejor desempeño global respecto al segundo grupo. Grantcharov y cols. (Grantcharov *et al.*, 2004) demostraron menor tiempo operatorio, tasa de error y mayor economía de movimientos en una colecistectomía laparoscópica en paciente real previo entrenamiento con simulación mediante realidad virtual respecto a un grupo control. Stefanidis y cols (Stefanidis *et al.*, 2013) encontraron que el 71% de residentes novatos de cirugía entrenados en un simulador mantenían sus habilidades en el EOR.

Respecto a los procedimientos endoscópicos, se ha visto que el entrenamiento mediante simulación previo a enfrentamiento al paciente en colonoscopia, sigmoidoscopia y endoscopia digestiva alta ha demostrado presentar ventajas sobre los grupos control no entrenados (Ahlberg *et al.*, 2005; Cohen *et al.*, 2006; Park *et al.*, 2007). Estos estudios mostraron que los residentes entrenados mediante simulación tuvieron un mejor desempeño en el paciente real que sus contrapartes en las etapas iniciales de la curva de aprendizaje, luego de lo cual no hubo diferencias significativas. Se ha visto también que los participantes sometidos a entrenamiento mediante simulación generan menos malestar en el paciente respecto a su contraparte (Sedlack & Kolars, 2004; Ahlberg *et al.*, 2005).

Se ha impulsado a lo largo de nuestro país el desarrollo de centros de simulación quirúrgica, apuntando al objetivo de expandir el acceso a entrenamientos en cirugía laparoscópica básica y avanzada en regiones, en consonancia con los proyectos recientemente propuestos por la Sociedad de Cirujanos de Chile (Jarufe & Barra, 2017). Sin embargo, este entrenamiento se ha visto obstaculizado por la falta de tutores expertos capaces de guiar las sesiones y entregar una adecuada retroalimentación a los alumnos. En este contexto, la Tele-simulación corresponde a un concepto que conjuga la supervisión a distancia (en inglés *tele-mentoring*) y la simulación, y que ha demostrado ser un método costo-efectivo de enseñanza en países de escasos recursos y puede ser utilizado tanto para evaluar como para enseñar habilidades laparoscópicas en cualquier lugar del mundo con acceso a internet (Okraimec *et al.*, 2010). Además, ha demostrado mejorar los conocimientos, la auto-confianza y el grado de comodidad en realizar procedimientos en países en desarrollo (Mikrogianakis *et al.*, 2011). Actualmente en nuestro país existen proyectos en marcha para implementar y validar herramientas de simulación a distancia, con plataformas virtuales que faciliten la evaluación y la entrega de la retroalimentación necesaria a alumnos en distintos centros de formación regionales de nuestro país.

Grabador de Datos (Caja Negra)

Durante la atención quirúrgica en el escenario operatorio, existen múltiples aspectos que pueden ser difíciles de notar o incluso llegar a pasarse por alto al evaluarse en tiempo real. La revisión de material audiovisual del procedimiento permite una evaluación más cuidadosa de éste, incluyendo la presencia de un número variado de evaluadores, utilizando distintas herramientas y escalas para obtener una apreciación completa de las competencias del operador (Goldenberg & Grantcharov, 2016).

El análisis de videos del escenario operatorio corresponde a una herramienta efectiva para mejorar la seguridad del paciente, sin embargo, ha sido poco utilizada respecto a otras áreas científicas, donde su uso se ha ampliado (Bezemer *et al.*, 2017). Un método sistemático de grabación y análisis de videos ha sido propuesto como estrategia para mejorar los cuidados en cirugía, prevenir complicaciones y permitir un efectivo aprendizaje de los errores (Gambadauro & Magos, 2012).

Los modelos tradicionales de análisis de videos en el escenario operatorio se basan, en su gran mayoría, en el análisis retrospectivo de la ejecución aislada del procedimiento mediante técnicas mínimamente invasivas, principalmente por la captura de video intracorporal necesaria para llevarse a cabo. Esta modalidad sólo entrega una visión parcial de lo acontecido en el acto quirúrgico, centrándose en el análisis de las habilidades técnicas del operador, sin correlacionar éstas con el impacto real, por ejemplo, en los signos vitales del paciente o posibles daños o perjuicios hacia éste.

Tradicionalmente, se conoce como “Caja Negra” al dispositivo encargado del almacenamiento de datos respecto a la actividad de los instrumentos y las conversaciones de la tripulación de una aeronave. Su principal función es el análisis de los datos que permitan obtener información sobre lo ocurrido en los momentos previos a un eventual accidente. Utilizando un principio similar e intentando dar respuesta a la problemática antes expuesta, surge la Caja Negra (*Black Box*) (Goldenberg *et al.*, 2017), que corresponde a una innovadora herramienta que contiene y analiza una gran cantidad de datos del pabellón quirúrgico, incluyendo material audiovisual, parámetros fisiológicos del paciente, factores ambientales y múltiples aparatos de grabación. El video es capturado utilizando cámaras de ángulo amplio y además registra datos intracorporales del paciente mediante cámaras laparoscópicas o robóticas, o incluso con cámaras montadas en cirugías abiertas. Estos datos son posteriormente almacenados en un servidor para su posterior análisis por expertos.

El análisis prospectivo que genera esta nueva herramienta permite detectar con mayor sensibilidad pequeñas variaciones en la

ejecución de los procedimientos que podrían llegar a significar un daño hacia el paciente. Otra de las ventajas que proporciona es que no requiere la presencia física del examinador en el pabellón para evaluar la intervención, sino que puede hacerlo desde un lugar separado, haciendo de este método una herramienta menos invasiva para el paciente y para el mismo operador.

En nuestro país existen actualmente proyectos en desarrollo para la incorporación de esta nueva herramienta en la evaluación tanto de residentes de especialidades quirúrgicas como de cirujanos ya formados.

Conclusiones

Numerosos métodos de evaluación de las habilidades técnicas en cirugía han sido desarrollados en pro de mejorar los programas existentes de formación de especialistas del área quirúrgicas. Gran parte de éstos han sido validados en distintos estudios para una amplia gama de procedimientos, sin embargo, su aplicación actual continúa siendo escasa y no estructurada. Las escalas de evaluación constituyen una herramienta validada, accesible y de fácil aplicación, por lo que su uso estructurado debiera generalizarse en las distintas actividades de los programas de formación en cirugía. Por otro lado, los modelos de simulación han cobrado una importancia excluyente en la preparación, evaluación y adiestramiento en el área quirúrgica. Nuevos desafíos como la implementación de herramientas de simulación a distancia y la incorporación de instrumentos complejos como la Caja Negra asoman como proyectos ambiciosos tendientes a optimizar las herramientas de evaluación y retroalimentación en la formación de los futuros cirujanos.

Referencias

- Aggarwal R, Grantcharov TP & Darzi A (2007). Framework for Systematic Training and Assessment of Technical Skills. *J Am Coll Surg* **204**, 697–705.
- Aggarwal R, Ward J, Balasundaram I, Sains P, Athanasiou T & Darzi A (2007). Proving the effectiveness of virtual reality simulation for training in laparoscopic surgery. *Ann Surg* **246**, 771–9.
- Ahlberg G, Hultcrantz R, Jaramillo E, Lindblom A & Arvidsson D (2005). Virtual reality colonoscopy simulation: a compulsory practice for the future colonoscopist? *Endoscopy* **37**, 1198–1204.
- Beard J, Marriott J, Purdie H & Crossley J (2011). Assessing the surgical skills of trainees in the operating theatre: a prospective observational study of the methodology. *Health Technol Assess* **15**.

- Bezemer J, Cope A, Korikiakangas T, Kress G, Murtagh G, Weldon S-M & Kneebone R (2017). Microanalysis of video from the operating room: an underused approach to patient safety research. *BMJ Qual Saf.* **26**, 583–7.
- Buckley CE, Kavanagh DO, Traynor O & Neary PC (2014). Is the skill-set obtained in surgical simulation transferable to the operating theatre? *Am J Surg.* **207**, 146–57.
- Cohen J, Cohen SA, Vora KC, Xue X, Burdick JS, Bank S, Bini EJ, Bodenheimer H, Cerulli M, Gerdes H, Greenwald D, Gress F, Grosman I, Hawes R, Mullin G, Schnoll-Sussman F, Starpoli A, Stevens P, Tenner S & Villanueva G (2006). Multicenter, randomized, controlled trial of virtual-reality simulator training in acquisition of competency in colonoscopy. *Gastrointest Endosc.* **64**, 361–368.
- Ende A, Zopf Y, Konturek P, Naegel A, Hahn EG, Matthes K & Mais J (2012). Strategies for training in diagnostic upper endoscopy: a prospective, randomized trial. *Gastrointest Endosc.* **75**, 254 – 260.
- Eubanks TR, Clements RH, Pohl D, Williams N, Schaad DC, Horgan S & Pellegrini C (1999). An objective scoring system for laparoscopic cholecystectomy. *J Am Coll Surg.* **189**, 566–574.
- Gallagher AG, Ritter EM & Satava RM (2003). Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. *Surg Endosc;* **17**, 1525–1529
- Gambadauro P & Magos A (2012). Surgical videos for accident analysis, performance improvement, and complication prevention: Time for a surgical black box? *Surg Innov.* **19**, 76–80.
- Garbens A, Goldenberg M, Wallis CJD, Tricco A & Grantcharov TP (2018). The cost of intraoperative adverse events in abdominal and pelvic surgery: A systematic review. *Am J Surg.* **215**, 163-170.
- Goldenberg MG & Grantcharov TP (2016). Video-analysis for the assessment of practical skill. *Tijdschr voor Urol.* **6**, 128–36.
- Goldenberg MG, Jung J & Grantcharov TP (2017). Using Data to Enhance Performance and Improve Quality and Safety in Surgery. *JAMA Surg.* **152**, 972.
- Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J & Funch-Jensen P (2004). Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg.* **91**, 146–50.
- Grantcharov TP & Reznick RK (2008). Teaching procedural skills. *Bmj.* **336**, 1129–31.
- Grantcharov TP, Schulze S & Kristiansen VB (2007). The impact of objective assessment and constructive feedback on improvement of laparoscopic performance in the operating room. *Surg Endosc Other Interv Tech.* **21**, 2240–3.
- Gumbs AA, Hogle NJ & Fowler DL (2007). Evaluation of Resident Laparoscopic Performance Using Global Operative Assessment of Laparoscopic Skills. *J Am Coll Surg.* **204**, 308–13.
- Hatala R, Cook DA, Brydges R & Hawkins R (2015). Constructing a validity argument for the Objective Structured Assessment of Technical Skills (OSATS): a systematic review of validity evidence. *Adv Heal Sci Educ.* **20**, 1149–75.
- Healey AN, Sevdalis N & Vincent CA (2006). Measuring intra-operative interference from distraction and interruption observed in the operating theatre. *Ergonomics.* **49**, 589–604.
- Hogle NJ, Liu Y, Ogden RT & Fowler DL (2014). Evaluation of surgical fellows' laparoscopic performance using Global Operative Assessment of Laparoscopic Skills (GOALS). *Surg Endosc Other Interv Tech.* **28**, 1284–90.
- Jarufe N & Barra M (2017). Comenzando un nuevo periodo con énfasis en la educación quirúrgica y las regiones. *Rev Chil Cir.* **69**, 103-104.
- Korndorffer JR, Dunne JB, Sierra R, Stefanidis D, Touchard CL & Scott DJ (2005). Simulator training for laparoscopic suturing using performance goals translates to the operating room. *J Am Coll Surg.* **201**, 23–9.
- Kurashima Y, Feldman LS, Al-Sabah S, Kaneva PA, Fried GM & Vassiliou MC (2011). A tool for training and evaluation of laparoscopic inguinal hernia repair: The global operative assessment of laparoscopic skills-groin hernia (GOALS-GH). *Am J Surg.* **201**, 54–61
- Martin JA, Regehr G, Reznick R, Macrae H, Murnaghan J, Hutchison C & Brown M (1997). Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg.* **84**, 273–8.
- Mikrogianakis A, Kam A, Silver S, Bakanisi B, Henao O, Okrainec A & Azzie G (2011). Telesimulation: An Innovative and Effective Tool for Teaching Novel Intraosseous Insertion Techniques in Developing Countries. *Acad Emerg Med.* **18**,420–7.
- Moya R P, Ruz A M, Parraguez L E, Carreño E V, Rodríguez C AM & Froes M P (2017). Efectividad de la simulación en la educación médica desde la perspectiva de seguridad de pacientes. *Rev Med Chil.* **145**, 514–26

- Niitsu H, Hirabayashi N, Yoshimitsu M, Mimura T, Taomoto J, Sugiyama Y, Murakami S, Saeki S, Mukaida H & Takiyama W (2013). Using the Objective Structured Assessment of Technical Skills (OSATS) global rating scale to evaluate the skills of surgical trainees in the operating room. *Surg Today* **43**, 271–5
- Okraïneç A, Henao O & Azzie G (2010). Telesimulation: an effective method for teaching the fundamentals of laparoscopic surgery in resource-restricted countries. *Surg Endosc*. **24**, 417–22
- Park J, Mac Rae H, Musselman LJ, Rossos P, Hamstra SJ, Wolman S & Reznick RK (2007). Randomized controlled trial of virtual reality simulator training: transfer to live patients. *Am J Surg* **194**, 205–211
- Sachs TE & Pawlik TM (2015). See one, do one, and teach none: Resident experience as a teaching assistant. *J Surg Res* **195**, 44–51
- Sarker SK, Chang A, Vincent C & Darzi SA (2006). Development of assessing generic and specific technical skills in laparoscopic surgery. *Am J Surg* **191**, 238–244.
- Shippey SH, Chen TL, Chou B, Knoepp LR, Bowen CW & Handa VL (2011). Teaching subcuticular suturing to medical students: video versus expert instructor feedback. *J Surg Educ* **68**, 397–402.
- Stefanidis D & Acker C (2008). Proficiency-Based Laparoscopic Simulator Training Leads to Improved Operating Room Skill That Is Resistant to Decay. *Surg Innov* **15**, 69–73.
- Stefanidis D, Korndorffer JR, Markley S, Sierra R, Heniford BT & Scott DJ (2007). Closing the Gap in Operative Performance Between Novices and Experts: Does Harder Mean Better for Laparoscopic Simulator Training? *J Am Coll Surg*. **205**, 307–13
- Stefanidis D, Yonce TC, Korndorffer JR Jr, Phillips R & Coker A (2013). Does the incorporation of motion metrics into the existing FLS metrics lead to improved skill acquisition on simulators? A single blinded, randomized controlled trial. *Ann Surg*. **258**, 46–52
- Vaillancourt M, Ghaderi I, Kaneva P, Vassiliou M, Kolozsvari N, George I, Sutton FE, Seagull FJ, Park AE, Fried GM & Feldman LS (2011). GOALS-incisional hernia: a valid assessment of simulated laparoscopic incisional hernia repair. *Surg Innov*. **18**, 48–54
- Van Hove PD, Tuijthof GJ, Verdaasdonk EG, Stassen LP & Dankelman J (2010). Objective assessment of technical surgical skills. *Br J Surg* **97**, 972–87
- Vassiliou MC, Feldman LS, Andrew CG, Bergman S, Leffondré K, Stanbridge D & Fried GM (2005). A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg*. **190**, 107–13
- Vassiliou MC, Kaneva PA, Poulouse BK, Dunkin BJ, Marks JM, Sadik R, Sroka G, Anvari M, Thaler K, Adrales GL, Hazey JW, Lightdale JR, Velanovich V, Swanstrom LL, Mellinger JD & Fried GM (2010). Global assessment of gastrointestinal endoscopic skills (GAGES): A valid measurement tool for technical skills in flexible endoscopy. *Surg Endosc Other Interv Tech*. **24**, 1834–41