



Original

Desarrollo y estudio piloto de un conjunto esencial de indicadores para los servicios de cirugía general[☆]



CrossMark

**Victor Soria-Aledo^{a,b}, Daniel Angel-Garcia^c, Ismael Martinez-Nicolas^{c,*},
Pere Rebasa Cladera^d, Roger Cabezali Sanchez^e y Luis Francisco Pereira García^f**

^a Sección de Gestión de Calidad de la Asociación Española de Cirujanos, Servicio de Cirugía General, Hospital Morales Meseguer, Murcia, España

^b Departamento de Cirugía, Facultad de Medicina, Universidad de Murcia, Murcia, España

^c Centro de Investigación en Evaluación y Encuestas, Instituto Nacional de Salud Pública, Cuernavaca, Morelos, México

^d Sección de Gestión de Calidad de la Asociación Española de Cirujanos, Departamento de Cirugía, Corporación Sanitaria Parc Taulí, Sabadell, Barcelona, España

^e Sección de Gestión de Calidad de la Asociación Española de Cirujanos, Servicio de Cirugía, Fundación Hospital de Calahorra, Calahorra, La Rioja, España

^f Hospital General Universitario Los Arcos del Mar Menor, Pozo Aledo, Murcia, España

INFORMACIÓN DEL ARTÍCULO

RESUMEN

Historia del artículo:

Recibido el 18 de noviembre de 2015

Aceptado el 27 de junio de 2016

On-line el 4 de agosto de 2016

Palabras clave:

Sistemas de información en salud

Indicadores de calidad de la atención de salud

Servicio de cirugía en hospital

Cirugía general

Benchmarking

Introducción: En la actualidad no se dispone de un conjunto adecuado de indicadores para benchmarking en las unidades de cirugía general del Sistema Nacional de Salud. Este trabajo presenta la selección, el desarrollo y los resultados del estudio piloto de un primer grupo de indicadores para esta finalidad.

Métodos: Se realizó una selección y priorización de indicadores mediante un Delphi modificado con un grupo de expertos de la Asociación Española de Cirujanos. Los indicadores priorizados fueron sometidos a un estudio cualitativo de factibilidad y, para aquellos medidos por historia clínica, cuál-cuantitativo de fiabilidad en un hospital público. Se obtuvieron resultados de concordancia simple y estadístico kappa, ajustado y no ajustado por prevalencias y sesgos, para 3 evaluadores con un muestreo aleatorio sistemático de 30 casos por indicador.

Resultados: De los 13 indicadores propuestos, 12 resultaron factibles (5 de historia clínica y 7 de bases de datos). De los 5 de historia, 3 resultaron fiables (concordancia interobservador > 95% o índice kappa > 0,6 para compuestos y subindicadores, o bien kappa ajustado por prevalencias y sesgos > 0,6 en presencia de prevalencias extremas) y 2 necesitaron ser redefinidos a partir de los resultados obtenidos.

[☆] Parte de esta publicación fue presentada como comunicación oral en el X Congreso Regional de Calidad Asistencial, celebrado entre los días 16 y 17 de junio de 2015, bajo el nombre de «Pilotaje de un set de indicadores de benchmarking en cirugía general».

* Autor para correspondencia.

Correos electrónicos: ismael.martinez@insp.mx, ismael.m.n@um.es (I. Martínez-Nicolas).

<http://dx.doi.org/10.1016/j.ciresp.2016.06.009>

0009-739X/© 2016 AEC. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Conclusiones: Los 5 indicadores de historia clínica podrán utilizarse para comparar unidades quirúrgicas, mientras que los 7 indicadores factibles de bases de datos necesitarán mayor validación y ajuste de riesgo para permitir comparaciones entre servicios. Los resultados del centro evaluado muestran áreas de mejora en algunos procesos de la atención.

© 2016 AEC. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Development and pilot study of an essential set of indicators for general surgery services

A B S T R A C T

Keywords:

Health information systems
Quality indicators
Surgery department
General surgery
Benchmarking

Introduction: At present there is a lack of appropriate quality measures for benchmarking in general surgery units of Spanish National Health System. The aim of this study is to present the selection, development and pilot-testing of an initial set of surgical quality indicators for this purpose.

Methods: A modified Delphi was performed with experts from the Spanish Surgeons Association in order to prioritize previously selected indicators. Then, a pilot study was carried out in a public hospital encompassing qualitative analysis of feasibility for prioritized indicators and an additional qualitative and quantitative three-rater reliability assessment for medical record-based indicators. Observed inter-rater agreement, prevalence adjusted and bias adjusted kappa and non-adjusted kappa were performed, using a systematic random sample ($n = 30$) for each of these indicators.

Results: Twelve out of 13 proposed indicators were feasible: 5 medical record-based indicators and 7 indicators based on administrative databases. From medical record-based indicators, 3 were reliable (observed agreement > 95%, adjusted kappa index > 0.6 or non-adjusted kappa index > 0.6 for composites and its components) and 2 needed further refinement.

Conclusions: Currently, medical record-based indicators could be used for comparison purposes, whilst further research must be done for validation and risk-adjustment of outcome indicators from administrative databases. Compliance results in the adequacy of informed consent, diagnosis-to-treatment delay in colorectal cancer, and antibiotic prophylaxis show room for improvement in the pilot-tested hospital.

© 2016 AEC. Published by Elsevier España, S.L.U. All rights reserved.

Introducción

La relación de los servicios de cirugía general (SCG) con la calidad no es nueva. Desde la seguridad del paciente^{1,2} a la mejora de procesos de atención³, se encuentran numerosos elementos de la gestión de calidad que los servicios de cirugía han adaptado rápidamente con el objetivo de mejorar su actividad profesional. El desarrollo de instrumentos para *benchmarking* y el aprendizaje comparativo se ha propuesto como motor de mejora dentro de la especialidad⁴. Sin embargo, a pesar de la existencia de indicadores que potencialmente podrían ser útiles para este fin⁵⁻⁷, en cirugía aún son escasas las experiencias de comparación entre servicios hospitalarios. Las únicas experiencias publicadas de *benchmarking* en SCG se circunscriben al ámbito de la cirugía mayor ambulatoria⁶ y la cirugía torácica⁸.

Actualmente no existe un conjunto equilibrado de indicadores que sirva para comparar y facilitar la mejora continua de la calidad en los SCG. Por ello, la Asociación Española de Cirujanos, a través de su Sección de Gestión de Calidad, ha tomado esta iniciativa. Con el objetivo último de crear un conjunto de indicadores para monitorización y *benchmarking*

en los SCG de los hospitales de agudos del Sistema Nacional de Salud (SNS), este trabajo presenta la selección, el desarrollo y los resultados del estudio piloto de un primer grupo de indicadores.

Métodos

En este estudio se siguieron 3 fases sucesivas: 1) identificación de indicadores de calidad asistencial y su priorización mediante métodos de consenso; 2) elaboración y adaptación de las fichas técnicas de los indicadores seleccionados, y 3) estudio piloto de los indicadores e instrumentos en casos del servicio de cirugía general y digestiva de un hospital público. Estas etapas se llevaron a cabo entre julio y diciembre de 2014.

Identificación del bloque inicial y priorización de los indicadores

Los objetivos de esta fase fueron la identificación de un grupo de indicadores que: a) incluyeran aspectos relativos a todo el proceso asistencial; b) representaran aspectos básicos que sirviesen para evaluar la calidad de la asistencia de un SCG de

cualquier hospital de agudos del SNS; c) tuvieran una clara evidencia científica o un fuerte acuerdo de expertos, y d) sirvieran para implementar medidas de mejora a partir de sus resultados.

Se realizó una búsqueda en la base de la Agency for Healthcare Research and Quality (AHRQ) y en las bases de datos Medline, EMBASE y The Cochrane Library desde 2000 hasta 2014, ambos incluidos, utilizando los términos: «Quality indicators» AND «General surgery»; y «Quality indicators» AND «Digestive surgery». Se estableció un método de consenso y priorización tipo RAND (Delphi modificado)⁹ en el que participaron un grupo de 8 cirujanos pertenecientes a la Sección de Gestión de Calidad de la Asociación Española de Cirujanos y 2 expertos en Calidad Asistencial de la Consejería de Sanidad y Política Social de la Región de Murcia. Se establecieron 3 rondas de consenso (2 mediante correo electrónico y una reunión presencial), sintetizadas en la figura 1.

Elaboración y adaptación de las fichas técnicas de los indicadores

Se solicitó una consultoría a 2 expertos externos para llevar a cabo la confección de los indicadores en un formato estandarizado (fig. 2). En este proceso, las fichas técnicas de indicadores originales fueron adaptadas, traducidas al español (en su caso) y trabajadas minuciosamente en sus

especificaciones. Para ello, los consultores redactaron varios borradores que fueron revisados de forma iterativa por expertos del grupo de trabajo, hasta que ya no hubo comentarios ni modificaciones al respecto.

Estudio piloto

Esta fase se desarrolló en el Hospital General Universitario Morales Meseguer de Murcia. Este es un hospital público con 340 camas y con una población adscrita de más de 280.000 habitantes.

Para proceder de forma operativa, se estudiaron los indicadores de forma separada según su fuente de datos.

Indicadores basados en historias clínicas

Se realizó una evaluación retrospectiva y externa por 3 evaluadores independientes. Dos evaluadores experimentados (Eva1 y Eva2) recogieron los datos de forma concomitante, mientras que un tercer evaluador (Eva3), un cirujano entrenado únicamente con los materiales generados para el estudio piloto —es decir, fichas técnicas y formatos de recogida de datos—, lo hizo 2 meses después.

Se consideró una muestra de 30 casos por indicador. Esta fue seleccionada por muestreo aleatorio sistemático del primer semestre de 2014. Para la identificación de casos se utilizaron listados de altas quirúrgicas extraídos del Conjunto

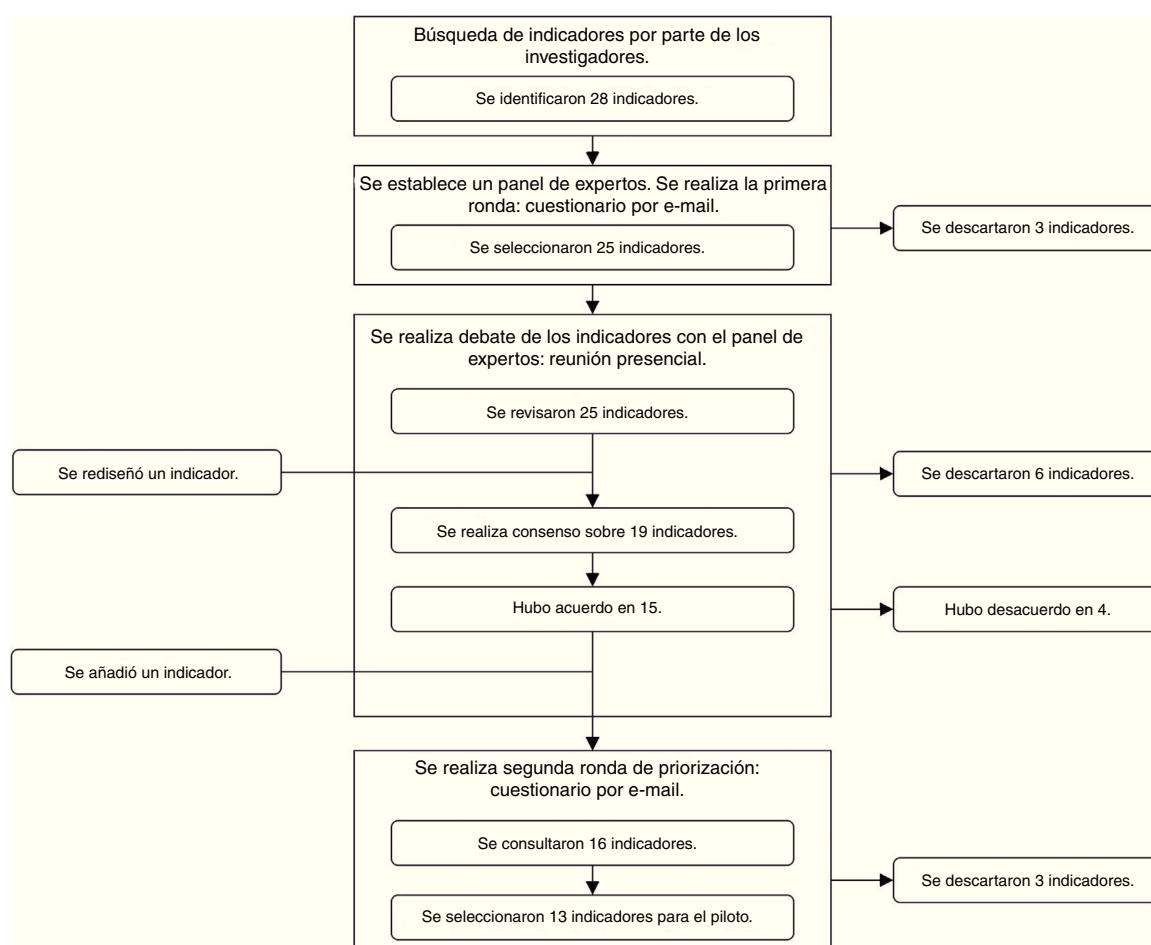


Figura 1 – Proceso de selección y priorización de indicadores.

(Nombre del indicador)
Set de indicadores: (Nombre del Set)
Subgrupo:
ID indicador:
Nombre abreviado: (Nombre abreviado como fórmula)
Descripción:
Justificación:
Tipo de medida: (Estructura/Proceso/Resultado)
Mejora identificada como: (Aumento/Disminución del resultado)
Numerador:
Las poblaciones incluidas:
Poblaciones excluidas:
Elementos de datos: (Datos utilizados para identificación de numerador)
Denominador:
Las poblaciones incluidas:
Poblaciones excluidas:
Elementos de datos: (Datos utilizados para identificación de numerador)
Ajuste de riesgo:
Enfoque de la recogida de datos: (Método de evaluación; Fuente de datos)
Precisión de los datos:
Sugerencias en el análisis de la medición:
Muestreo: (Tipo de muestreo; Tamaño muestral sugerido)
Datos reportados como: (Tipo de datos estadísticos)
Origen del indicador:
Referencias seleccionadas:
Notas: (Notas u observaciones para la correcta evaluación del indicador)

Figura 2 – Formato estandarizado de indicador.

Mínimo Básico de Datos (CMBD) usando los criterios descritos en la ficha técnica de cada indicador¹⁰.

Se exploraron 3 atributos de este tipo de indicadores: factibilidad, fiabilidad y utilidad para la identificación de oportunidades de mejora a partir de la estimación de cumplimiento.

Factibilidad de medición. Se consideró factible todo aquel indicador que pudo ser evaluado con la información disponible en el hospital, alcanzando el tamaño muestral establecido ($n = 30$). Adicionalmente, se realizó un análisis cualitativo con información recogida durante la evaluación.

Fiabilidad interobservador. Para valorar la fiabilidad de los indicadores se siguieron los siguientes pasos: 1) analizar el índice de concordancia general (P_o) de los 3 evaluadores —porcentaje de casos con la misma respuesta—, debiendo ser $\geq 95\%$ para aceptar como fiable el indicador; 2) interpretar la fiabilidad a través del índice kappa (κ) para 3 evaluadores¹¹: solo se tuvieron en cuenta como aceptables valores de kappa $> 0,6$ según los criterios de Landis y Koch¹², y 3) indagar si la prevalencia de la característica evaluada era extrema (cumplimiento o incumplimiento $\geq 85\%$) como posible origen de desviación de κ hacia 0, lo que no permite interpretar los resultados de forma fidedigna¹³.

Para enfrentar este problema, así como para observar diferencias entre una medición realizada por evaluadores experimentados y un cirujano entrenado, se analizaron los supuestos anteriores por cada pareja de evaluadores y se añadió el cálculo del índice kappa ajustado por sesgos y prevalencias (PABAK, por sus siglas en inglés)¹⁴, cuyos valores se interpretaron igual que κ .

Finalmente, se realizó un análisis cualitativo con información recabada durante la evaluación y la consulta a los evaluadores para obtener un mejor conocimiento de los problemas de fiabilidad existentes.

Utilidad para la identificación de oportunidades de mejora. Se obtuvo una estimación puntual del nivel de cumplimiento de todos los indicadores y subindicadores, así como su intervalo de confianza del 95%, utilizando las fórmulas para muestreo aleatorio sistemático no estratificado. Al ser los indicadores compuestos de tipo 100% de cumplimiento¹⁵ y de carácter binomial, se aplicaron las mismas fórmulas de estimación que en indicadores sencillos y subindicadores.

Indicadores basados en Conjunto Mínimo Básico de Datos
Se realizó una explotación de los registros oficiales del hospital, tomando los egresos quirúrgicos pertenecientes al año 2013. Para ello, se solicitó al servicio de documentación

una exportación del CMBD (tanto de atención hospitalaria como ambulatoria especializada) con variables descritas en un manual metodológico¹⁰. Las definiciones de estas variables pueden ser consultadas en el portal estadístico del Ministerio de Sanidad, Servicios Sociales e Igualdad^{16,17}.

Con estos datos, se exploró la factibilidad y la utilidad de estos indicadores.

Factibilidad de medición. Se consideró factible todo aquel indicador que pudo ser calculado con la información disponible en el hospital, sin requerir una base de datos externa. El método de cálculo de cada uno se explica aparte¹⁰.

Utilidad para la identificación de oportunidades de mejora. Se obtuvo una prevalencia de los indicadores a partir su cálculo. Los cálculos fueron realizados por un evaluador y revisados posteriormente por otro para minimizar errores en el tratamiento de la información.

Análisis estadístico

Todos los análisis se realizaron con hojas de cálculo programadas, Stata 13¹⁸, y EpiDat 4.1¹⁹.

Resultados

Identificación del conjunto inicial y priorización de los indicadores

En la búsqueda inicial se identificaron 28 indicadores. Tras las 3 rondas de priorización, se descartaron 16 indicadores y se añadió un indicador (fig. 1). Finalmente, 13 indicadores fueron propuestos la prueba piloto (tabla 1): 5 indicadores de revisión de historias y 8 indicadores de bases de datos electrónicas (CMBD).

Elaboración y adaptación de las fichas técnicas de los indicadores

Los 13 indicadores fueron creados desarrollando todas las especificaciones técnicas necesarias a partir de indicadores existentes o de la evidencia disponible. De ellos, 4 fueron adaptados de los Patient Safety Indicators de AHRQ²⁰, de los National Hospital Inpatient Quality Measures de la Joint Commission²¹ o de los Indicadores Clave del SNS²²; 5 fueron modificados de los propuestos en un trabajo previo de la Sociedad Española de Calidad Asistencial⁵, y 3 fueron de creación propia a partir de diversas fuentes. Las fichas técnicas, junto con la bibliografía y el origen de cada indicador, se encuentran accesibles en la página web de la Asociación Española de Cirujanos¹⁰.

Estudio piloto

Indicadores basados en historias clínicas

Factibilidad de la medición. Se revisaron un total de 144 historias hasta completar 30 unidades de estudio en cada uno de los 5 indicadores. Las dificultades encontradas en la evaluación fueron: 2.2 (Intervalo Diagnóstico-Terapéutico en

Cáncer Colorrectal), gran cantidad de exclusiones; 3.1 (Profilaxis Antitrombótica), identificación de prescripción de heparina de difícil localización en los registros durante los días de estancia en UCI/Reanimación o el día de alta; 3.2 (Profilaxis Antibiótica), diversas fuentes de datos: hora de administración, hora de inicio de cirugía, antibiótico y dosis en diferentes registros.

Fiabilidad interobservador. Los resultados del análisis cuantitativo de la fiabilidad se muestran en la tabla 2. Los indicadores de intervalo diagnóstico-terapéutico, tanto en cáncer de colon como de mama, se han mostrado fiables con un $\kappa > 0,6$. El indicador sobre documento de consentimiento informado (DCI) se evaluó con PABAK (debido a prevalencias extremas), y también resultó satisfactorio en todos sus componentes.

Con respecto a los indicadores de profilaxis, se obtuvieron κ menores a 0,6 (sin prevalencia extrema), por lo que no se puede concluir que estos indicadores sean fiables. Estos indicadores se ven afectados por la suma de pequeñas discordancias de todos sus subindicadores.

El análisis cualitativo mostró: a) la dificultad de realizar un adecuado recuento de los días de prescripción de anti-trombótico, así como la interpretación del medicamento y dosis adecuada en todos los días de ingreso, y b) que los registros pueden no ser lo suficientemente buenos para determinar con exactitud el tiempo real de administración de antibiótico antes de la operación.

Estimación del cumplimiento. La tabla 3 muestra la estimación puntual de estos indicadores. El indicador compuesto sobre el DCI obtuvo resultados muy bajos. Se encontró que el DCI estaba presente y era el correspondiente a la cirugía (93,33 y 100%, respectivamente), pero el subindicador relacionado con los datos del médico (3,33%) es el que provocó este bajo cumplimiento.

Indicadores basados en Conjunto Mínimo Básico de Datos

Factibilidad de la medición. Todos los indicadores pudieron ser calculados a partir del CMBD. Sin embargo, el indicador 4.7 (Reintervención después de cirugía programada) necesitó ser adaptado de «reintervención» a «reingreso» al no poder identificar pertinenteamente las reintervenciones durante un mismo episodio asistencial, y el indicador 4.2 (Reingreso en menos de 7 días tras Cirugía Mayor Ambulatoria) requirió información adicional al CMBD de atención hospitalaria (CMBD-AH) proveniente del de atención ambulatoria especializada (CMBD-AAE).

Prevalencia de eventos quirúrgicos. Los resultados obtenidos se muestran en la tabla 4 en forma de proporción.

Discusión

Este trabajo expone la selección y el estudio piloto de un primer conjunto de indicadores para los SCG de un proyecto que se pretende escalar posteriormente. Se ha conseguido seleccionar un conjunto de 13 indicadores, los cuales pudieron ser medidos con las fuentes de datos disponibles. Solo un indicador de CMBD (Reintervención hasta 30 días tras cirugía programada) tiene serias limitaciones en su medición y puede considerarse no factible. De los 5 indicadores de historia

Tabla 1 – Conjunto de indicadores propuesto para servicios de cirugía general

Indicadores propuestos	Numerador	Denominador	Fuente de datos
1.1 Correcta cumplimentación del documento de consentimiento informado	Pacientes a los que ha intervenido quirúrgicamente el servicio de cirugía general que tienen el documento de consentimiento informado correctamente cumplimentado	Pacientes a los que se ha intervenido quirúrgicamente en el servicio de cirugía general	Historia clínica
2.1 Estancia preoperatoria menor de un día	Pacientes intervenidos el mismo día o al día siguiente del ingreso (estancia preoperatoria ≤ 1 día)	Pacientes ingresados para cirugía programada	CMBD
2.2 Intervalo diagnóstico-terapéutico en cáncer colorrectal	Pacientes con cáncer colorrectal cuyo intervalo entre confirmación diagnóstica y tratamiento quirúrgico es ≤ 28 días naturales	Pacientes diagnosticados con cáncer colorrectal que han recibido tratamiento quirúrgico	Historia clínica
2.3 Intervalo diagnóstico-terapéutico en cáncer de mama	Pacientes con cáncer de mama cuyo intervalo entre confirmación diagnóstica y primer tratamiento quirúrgico es ≤ 4 semanas o 28 días naturales	Pacientes diagnosticadas de cáncer de mama cuyo primer tratamiento es quirúrgico	Historia clínica
3.1 Profilaxis tromboembólica en cáncer colorrectal	Pacientes intervenidos por carcinoma colorrectal a los que se practica resección de colon y/o recto con profilaxis tromboembólica correcta ^a	Pacientes intervenidos por carcinoma colorrectal con resección de colon y/o recto	Historia clínica
3.2 Profilaxis antibiótica en cáncer colorrectal	Pacientes intervenidos por carcinoma colorrectal a los que se practica resección de colon y/o recto con profilaxis antibiótica correcta ^a	Pacientes intervenidos por carcinoma colorrectal con resección de colon y/o recto	Historia clínica
4.1 Reingreso en menos de 30 días tras intervención quirúrgica	Pacientes que reingresan de forma no programada en un intervalo de 30 días tras el alta hospitalaria	Pacientes sometidos a intervención quirúrgica programada o urgente	CMBD
4.2 Reingreso en menos de 7 días tras cirugía mayor ambulatoria	Pacientes que ingresan en los 7 días posteriores a una intervención de CMA	Pacientes con intervenciones quirúrgicas programadas realizadas en los GRD de CMA de cirugía general más frecuentes	CMBD
4.3 Hemorragia o hematoma postoperatorios (reintervención)	Casos de hemorragia o hematoma que requieren procedimiento quirúrgico de control de la hemorragia	Altas quirúrgicas con código de procedimiento quirúrgico. Altas en GRD quirúrgicos que tienen fecha de intervención	CMBD
4.4 Dehiscencia abdominopélvica tras cirugía	Altas de pacientes > 14 años con código CIE-9 de cierre de disrupción de la pared abdominal	Altas de cirugía abdominopélvica en pacientes > 14 años	CMBD
4.5 Infección en cirugía limpia	Altas con infección de sitio quirúrgico en cualquier campo de diagnóstico secundario	Altas programadas con cirugía limpia (cirugía mama, cirugía hernia, cirugía tiroides y paratiroides)	CMBD
4.6 Infección en cirugía colorrectal	Altas con infección de sitio quirúrgico en cualquier campo de diagnóstico secundario	Altas programadas con cirugía colorrectal	CMBD
4.7 Reintervención después de cirugía programada	Pacientes que son llevados a quirófano de manera no programada durante el mismo ingreso o hasta 30 días desde la intervención previa	Total de pacientes intervenidos de forma programada	CMBD

CIE: Clasificación Internacional de Enfermedades; CMA: Cirugía Mayor Ambulatoria; CMBD: Conjunto Mínimo Básico de Datos; GRD: Grupos Relacionados por el Diagnóstico.

^a La definición de profilaxis «correcta» se encuentra en la ficha técnica de cada indicador¹⁰.

clínica, 3 pueden considerarse fiables, mientras que los 2 indicadores de profilaxis antibiótica y antitrombótica fueron rediseñados con los resultados de este estudio.

En los últimos años se ha incrementado el interés por la medida y mejora de la calidad en los SCG, y una de las herramientas para llevarlo a cabo es el *benchmarking*²³. Trabajos previos demuestran su utilidad en el campo de la cirugía^{4,24,25}. Sin embargo, en España todavía no hay un buen sistema de indicadores que pueda servir para ello. La revisión de

la literatura existente y los métodos de consenso, aceptados y utilizados ampliamente dentro de la comunidad científica²⁶⁻²⁸, han permitido seleccionar indicadores de gran relevancia para los SCG.

Este trabajo está en línea con investigaciones y estudios previos. Así, nuestro paquete de indicadores incluye, entre otros, 5 indicadores de una priorización de la Sociedad Española de Calidad Asistencial⁵, en la que se propusieron indicadores que no han sido utilizados en pruebas piloto

Tabla 2 – Concordancia general, índice kappa, PABAK y prevalencia de cada indicador para los 3 evaluadores o por parejas

Indicador	3 evaluadores			Eva1-Eva2			Eva1-Eva3			Eva2-Eva3		
	P _o	κ (IC 95%) ^a	P% ^b	PABAK	κ	P% ^b	PABAK	κ	P% ^b	PABAK	κ	P% ^b
1.1. DCI	93%	0,58	94%	0,87	0,47	93%	0,87	0,47	93%	1	1	97%
Está presente en la historia	97%	0,74	94%	0,93	0,65	95%	0,93	0,65	95%	1	1	97%
DCI corresponde con la intervención	100%	—	100%	1	—	100%	1	—	100%	1	—	100%
Compuesto 3 criterios:	93%	0,57	94%	0,86	0,47	93%	0,86	0,47	93%	1	1	96%
• Firma del prof. y n.º de colegiado	93%	0,57	94%	0,86	0,47	93%	0,86	0,47	93%	1	1	96%
• Firma del paciente o representante	89%	-0,04	96%	0,93	0	98%	0,86	0	96%	0,79	-0,05	95%
• Lugar, fecha	100%	—	100%	1	—	100%	1	—	100%	1	—	100%
2.2. Intervalo diagnóstico-terapéutico colon	83%	0,77 (0,57-0,96)	60%	1	1	57%	0,67	0,65	62%	0,67	0,65	62%
2.3. Intervalo diagnóstico-terapéutico mama	90%	0,63	90%	0,93	0,78	92%	0,8	0,52	88%	0,87	0,63	90%
3.1. Profilaxis antitrombótica	73%	0,56 (0,29-0,82)	72%	0,8	0,78	65%	0,67	0,53	78%	0,47	0,37	73%
Prescripción diaria ingreso	77%	0,14	90%	0,67	0,36	85%	0,8	0	95%	0,6	0	90%
Medicamento adecuado	100%	—	100%	1	—	100%	1	—	100%	1	—	100%
Dosis adecuada	90%	0,53	92%	0,87	0,63	90%	0,93	0,65	95%	0,8	0,37	92%
Prescripción alta +30 días	93%	0,81	86%	0,93	0,87	85%	0,87	0,71	87%	0,93	0,84	88%
3.2. Profilaxis antibiótica	70%	0,56 (0,31-0,81)	64%	0,75	0,72	60%	0,4	0,36	65%	0,67	0,62	68%
Medicamento adecuado	93%	0,48	96%	1	1	93%	0,87	0	97%	0,87	0	97%
Dosis adecuada	96%	-0,04	99%	0,93	0	98%	0,93	0	98%	1	—	100%
Administración previa (1 h)	77%	0,64 (0,40-0,89)	68%	0,8	0,78	65%	0,53	0,47	68%	0,73	0,68	70%

Eva1: evaluador 1; Eva2: evaluador 2; Eva 3: evaluador 3; κ, índice kappa; P_o: índice de concordancia general o acuerdo observado; P%: prevalencia; en negrita, valores considerados como buena fiabilidad.

^a Kappa e intervalo de confianza jackknife al 95% para 3 evaluadores. El intervalo de confianza solo se muestra si no hay prevalencia extrema.

^b Prevalencia de la característica evaluada más frecuente (ya sea cumplimiento o incumplimiento).

Tabla 3 – Cumplimiento estimado de los indicadores y subindicadores de revisión de historias (1.^{er} semestre 2014)

Indicador	Estimación ^a	IC 95%
1.1 DCI	3	±5
Presente	93	±7
DCI correcto	100	±0
3 criterios	3	±5
Médico	3	±5
Paciente o representante	93	±7
Lugar y fecha	100	±0
2.2 Intervalo D-T cáncer colorrectal	57	±15
2.3 Intervalo D-T cáncer de mama	97	±9
3.1 Profilaxis tromboembólica	63	±10
Prescrito todo el ingreso	77	±9
Medicamento adecuado	100	±0
Dosis adecuada	87	±7
Prescripción al alta	83	±8
3.2 Profilaxis antibiótica	57	±11
Medicamento adecuado	93	±5
Dosis adecuada	90	±6
Administración previa (1 h)	63	±10

DCI: documento de consentimiento informado; D-T: diagnóstico-terapéutico; IC: intervalo de confianza.

^a Estimación puntual de porcentaje de cumplimiento con intervalo de confianza para proporciones con muestreo aleatorio sistemático no estratificado.

ni evaluados hasta ahora. Dada la innegable necesidad de probar empíricamente cualquier indicador antes de su utilización²⁹⁻³¹, hemos realizado un estudio piloto siguiendo la metodología utilizada en anteriores proyectos³², incorporando un mayor análisis de la fiabilidad interobservador.

Un aspecto imprescindible para los indicadores con fin comparativo es la fiabilidad y consistencia de su medición. Por esto se han presentado los resultados de kappa y PABAK, recomendado por otros autores cuando existen prevalencias extremas y sobre la base de que un solo índice de concordancia es insuficiente^{14,33,34}. Los indicadores compuestos de profilaxis han resultado los más conflictivos, estando condicionados por sus subindicadores y la acumulación de sus errores. Una limitación podría ser el tamaño muestral utilizado. No existe consenso suficiente sobre este tema, aunque algunos autores recomiendan no menos de 30 observaciones³⁵.

Por otro lado, los resultados de cumplimiento muestran oportunidades de mejora en el centro piloto, lo que fundamenta la utilidad de estos indicadores. Así, la inadecuada cumplimentación del DCI arroja una nueva perspectiva a los estudios realizados en el entorno³⁶⁻³⁸, ya que estos se centran en el aspecto formal y no en el uso en sí de estos documentos. También el bajo número de casos de cáncer colorrectal que cumplen con el intervalo diagnóstico-terapéutico recomendado (57%; IC 95% ± 15%) parece estar en consonancia con resultados de la Estrategia en Cáncer del SNS³⁹ en el año 2009 (54,7%; IC 95%: 51,9-57,4%), mientras que en cáncer de mama

Tabla 4 – Prevalencia en indicadores de CMBD para altas del servicio de cirugía del Hospital General Universitario Morales Meseguer (año 2013)

Indicador	Eventos	Denominador	Prevalencia
2.1 Estancia preoperatoria ≤ 1 día	993	1.000	99,3%
4.1 Reingreso < 30 días tras intervención quirúrgica	95	1.481	6,4%
4.2 Reingreso < 7 días tras CMA	5	365	1,4%
4.3 Hemorragia/hematoma post-operatorios (reintervención)	40	1.000	4%
4.4 Dehiscencia abdominopélvica	8	454	1,8%
4.5 Infección cirugía limpia	2	414	0,5%
4.6 Infección cirugía colorrectal	11	135	8,1%
4.7 Reintervención en cirugía programada	4	927	0,4%

son mucho mejores (97%, IC 95% ± 9 frente a 43,6%, IC 95%: 41-46%). Sin embargo, los resultados de los indicadores de CMBD necesitan una comparación adecuada con un estándar de referencia, ya sea externo o del mismo hospital, para poder determinar áreas de mejora.

Para que esta comparación pueda ser válida en indicadores de resultado es necesario que se analice la precisión de los indicadores, así como que sean ajustados por el case-mix de la población atendida^{40,41}. Por otra parte, se ha mostrado factible construir y medir los indicadores de CMBD previstos, lo que permite continuar la investigación sobre la precisión de estos, así como sobre el ajuste necesario para su comparación entre centros.

En conclusión, el conjunto de 13 indicadores orientado al benchmarking de los SCG que se presenta en este trabajo, salvo un indicador, ha demostrado ser factible. Dos de los indicadores de historias han sido revisados y re-evaluados para asegurar su validez. Se requiere más investigación en el ajuste del riesgo de los indicadores de resultado, así como en la computación y automatización de los indicadores de CMBD. Cinco indicadores pueden ser usados actualmente para buscar oportunidades de mejora y hacer comparaciones entre centros sin necesidad de ningún ajuste.

Financiación

Financiado por la Asociación Española de Cirujanos: «Proyecto elaboración de indicadores de la AEC».

Autoría

Victor Soria-Aledo contribuyó en la concepción y diseño del estudio, en la escritura del artículo y en la aprobación final para su publicación.

Daniel Angel-García contribuyó en la recogida, análisis e interpretación de los datos, redacción y revisión crítica del artículo, y aprobación final para su publicación.

Ismael Martínez-Nicolás contribuyó en el diseño del estudio y del trabajo de campo, en la recogida, análisis e interpretación de los datos, en la escritura del artículo y en la aprobación final para su publicación.

Pere Rebasa Cladera contribuyó en la concepción inicial y diseño del estudio, en la participación y coordinación del grupo de expertos, en la revisión crítica con importantes contribuciones intelectuales y en la aprobación de la versión final para su publicación.

Roger Cabezali Sanchez contribuyó en la concepción inicial, en la participación del grupo de expertos, en la redacción, revisión crítica y aprobación de la versión final para su publicación.

Luis Francisco Pereira García contribuyó en la recogida y análisis de los datos, en la redacción, revisión crítica y aprobación de la versión final para su publicación.

Conflictos de intereses

Los autores declaran no tener ningún conflicto de interés laboral, de investigación, económico ni moral durante la realización de este trabajo. Las conclusiones que de él se derivan son absolutamente independientes y no hay ninguna relación con agentes externos que pueda haber influido en estas.

Agradecimientos

Al Programa EMCA de la Consejería de Sanidad de la Región de Murcia. A Pedro Parra Hidalgo y José Eduardo Calle Urra, expertos en calidad asistencial de la Región de Murcia, por su asesoría y colaboración durante el desarrollo de indicadores. A la Asociación Española de Cirujanos por su involucración, apoyo y financiación. Y a todos los expertos que participaron en la fase de selección y priorización de indicadores.

BIBLIOGRAFÍA

1. De Vries EN, Prins HA, Crolla RMPH, den Outer AJ, van Andel G, van Helden SH, et al. Effect of a comprehensive surgical safety system on patient outcomes. *N Engl J Med.* 2010;363:1928-37.
2. Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat A-HS, Dellinger EP, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med.* 2009;360:491-9.
3. Cevasco M, Ashley SW. Quality measurement and improvement in general surgery. *Perm J.* 2011;15:48-53.
4. Hernández-Navarrete MJ, Arribas-Llorente JL, Solano-Bernad VM, Misiego-Peral A, Rodríguez-García J, Fernández-García JL, et al. Programa de mejora de la infección nosocomial en cirugía por cáncer colorrectal [Quality improvement program of nosocomial infection in colorectal cancer surgery]. *Med Clin.* 2005;125:521-4.
5. Sociedad Española de Calidad Asistencial. Indicadores de calidad para hospitales del Sistema Nacional de Salud.

- Internet. 2012. 111 pp. [consultado 14 Feb 2016]. Disponible en: <http://www.calidadasistencial.es/>
6. Martínez Rodenas F, Codina Grifell J, Deulofeu Quintana P, Garrido Corchón J, Blasco Casares F, Gibanel Garanto X, et al. Indicadores de calidad asistencial en cirugía mayor ambulatoria (2010-2012) [Indicators of healthcare quality in day surgery (2010-2012)]. Rev Calid Asist. 2014;29:172-9.
 7. Jiménez A, Elia M, Antonio Gracia J, Artigas C, Lamata F, Martínez M. Indicadores de calidad asistencial en cirugía mayor ambulatoria [Clinical indicators to evaluate quality in ambulatory surgery]. Cir Esp. 2004;76:325-30.
 8. Varela G, Molins L, Astudillo J, María Borro J, Canalís E, Freixinet J, et al. Experiencia piloto de benchmarking en cirugía torácica: comparación de la casuística e indicadores de calidad en resección pulmonar [Pilot benchmarking study of thoracic surgery in Spain: Comparison of cases of lung resection and indicators of quality]. Arch Bronconeumol. 2006;42:267-72.
 9. Fink A, Kosecoff J, Chassin M, Brook RH. Consensus methods: Characteristics and guidelines for use. Am J Public Health. 1984;74:979-83.
 10. Asociación Española de Cirujanos (AEC). Set de Indicadores de Calidad de los Servicios de Cirugía General. Internet. Blog de la Asociación Española de Cirujanos. 2016 [consultado 5 Jun 2016]. Disponible en: <http://www.aecirujanos.es/wpaec/actualizacion-disponible-set-indicadores-calidad-los-servicios-cirugia-general/>
 11. Fleiss J. The Measurement of Interrater Agreement. Statistical Methods for Rates and Proportions. Hoboken, NJ: John Wiley & Sons, Inc.; 1982: 598-626.
 12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-74.
 13. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol. 1990;43: 543-9.
 14. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol. 1993;46:423-9.
 15. Saturno PJ. La invasión de los indicadores compuestos. Riesgos y beneficios para la gestión de la calidad. Rev Calid Asist. 2004;19:407-15.
 16. Ministerio de Sanidad Servicios Sociales e Igualdad (MSSSI). Registro de Altas de Hospitalización: CMBD. Sistema Nacional de Salud (Manual de definiciones y glosario de términos). Internet. Madrid; 2014 [consultado 26 Mar 2016]. Disponible en: <http://pestadistico.inteligenciadegestion.msssi.es>
 17. Ministerio de Sanidad Servicios Sociales e Igualdad (MSSSI). Registro de Altas: CMBD estatal. Atención Ambulatoria Especializada CMBD-AAE. Sistema Nacional de Salud (Manual de definiciones y glosario de términos). Internet. Madrid; 2014 [consultado 26 Mar 2016]. Disponible en: <http://pestadistico.inteligenciadegestion.msssi.es/>
 18. StataCorp. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP. 2013.
 19. Consellería de Sanidade, Xunta de Galicia, España; Organización Panamericana de la Salud (OPS-OMS); Universidad CES, Colombia. Epidat: programa para análisis epidemiológico de datos. 2014 [consultado 14 Jul 2015]. Disponible en: <http://dxsp.sergas.es/>
 20. Agency for Healthcare Research and Quality (AHRQ). Patient Safety Indicators Technical Specifications Updates — Version 5.0, March 2015. Internet. AHRQ Quality Indicators. 2015 [consultado 26 Mar 2016]. Disponible en: http://www.qualityindicators.ahrq.gov/Modules/PSI_TechSpec.aspx
 21. The Joint Commission. Specifications Manual for National Hospital Inpatient Quality Measures. Internet. The Joint Commission: Performance Measurement. 2015 [consultado 26 Mar 2016]. Disponible en: http://www.jointcommission.org/specifications_manual_for_national_hospital_inpatient_quality_measures.aspx
 22. Ministerio de Sanidad Servicios Sociales e Igualdad (MSSSI). Indicadores Clave Sistema Nacional de Salud. Versión 1.0. Internet. 2016 [consultado 26 Mar 2016]. Disponible en: <http://inclasns.msssi.es/>
 23. Mace SE. Continuous quality improvement for the clinical decision unit. J Healthc Qual. 2004;26:29-36.
 24. Meissner W, Ullrich K, Zwacka S. Benchmarking as a tool of continuous quality improvement in postoperative pain management. Eur J Anaesthesiol. 2006;23:142-8.
 25. Alderman AK, Collins ED, Streu R, Grotting JC, Sulkin AL, Neligan P, et al. Benchmarking outcomes in plastic surgery: National complication rates for abdominoplasty and breast augmentation. Plast Reconstr Surg. 2009;124:2127-33.
 26. Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: A systematic review. PLoS One. 2011;6:e20476.
 27. Shekelle PG. Quality indicators and performance measures: Methods for development need more standardization. J Clin Epidemiol. 2013;66:1338-9.
 28. Bobrovitz N, Parrilla JS, Santana M, Straus SE, Stelfox HT. A qualitative analysis of a consensus process to develop quality indicators of injury care. Implement Sci. 2013;8:45.
 29. Hofer TP, Bernstein SJ, Hayward RA, DeMonner S. Validating quality indicators for hospital care. Jt Comm J Qual Improv. 1997;23:455-67.
 30. Rubin HR, Pronovost P, Diette GB. From a process of care to a measure: The development and testing of a quality indicator. Int J Qual Health Care. 2001;13:489-96.
 31. Mainz J. Developing evidence-based clinical indicators: A state of the art methods primer. Int J Qual Health Care. 2003;15:i5-11.
 32. Saturno P, Martínez-Nicolas I, Robles-García I, López-Soriano F, Angel-García D. Development and pilot test of a new set of good practice indicators for chronic cancer pain management. Eur J Pain. 2015;19:28-38.
 33. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. J Clin Epidemiol. 1990;43:551-8.
 34. Lantz CA, Nebenzahl E. Behavior and interpretation of the kappa statistic: Resolution of the two paradoxes. J Clin Epidemiol. 1996;49:431-4.
 35. McHugh ML. Interrater reliability: The kappa statistic. Biochem Med (Zagreb). 2012;22:276-82.
 36. López-Picazo Ferrer JJ, Tomás García N. Evaluación y mejora de la comprensión de los documentos de consentimiento informado [Evaluation and improvement of the comprehension of informed consent documents]. Cir Esp. 2016;94:221-6.
 37. López-Picazo JJ, Tomás-García N, Calle-Urra JE, Parra-Hidalgo P, Valverde-Iniesta JJ. Introducción de un sistema de acreditación para los documentos de consentimiento informado hospitalario [Introduction of an accreditation system for hospital informed consent forms]. Rev Calid Asist. 2015;30:55-63.
 38. Calle-Urra JE, Parra-Hidalgo P, Saturno-Hernández PJ, Martínez-Martínez MJ, Navarro-Moya FJ. Evaluación de la calidad formal de los documentos de consentimiento informado en 9 hospitales [Formal quality assessment of informed consent documents in 9 hospitals]. Rev Calid Asist. 2013;28:234-43.
 39. Ministerio de Sanidad Servicios Sociales e Igualdad. Evaluación de la práctica asistencial oncológica. Estrategia en Cáncer del Sistema Nacional de Salud. Madrid: MSSSI. 2013.
 40. Mainz J. Defining and classifying clinical indicators for quality improvement. Int J Qual Health Care. 2003;15: 523-30.
 41. Iezzoni LI. The risks of risk adjustment. JAMA. 1997;278:1600-7.