



# CIRUGÍA ESPAÑOLA

www.elsevier.es/cirugia



## Original article

# Development and Pilot Study of an Essential Set of Indicators for General Surgery Services<sup>☆,☆☆</sup>



Victor Soria-Aledo,<sup>a,b</sup> Daniel Angel-Garcia,<sup>c</sup> Ismael Martinez-Nicolas,<sup>c,\*</sup>  
Pere Rebas Cladera,<sup>d</sup> Roger Cabezali Sanchez,<sup>e</sup> Luis Francisco Pereira García<sup>f</sup>

<sup>a</sup> Sección de Gestión de Calidad de la Asociación Española de Cirujanos, Servicio de Cirugía General, Hospital Morales Meseguer, Murcia, Spain

<sup>b</sup> Departamento de Cirugía, Facultad de Medicina, Universidad de Murcia, Murcia, Spain

<sup>c</sup> Centro de Investigación en Evaluación y Encuestas, Instituto Nacional de Salud Pública, Cuernavaca, Morelos, Mexico

<sup>d</sup> Sección de Gestión de Calidad de la Asociación Española de Cirujanos, Departamento de Cirugía, Corporación Sanitaria Parc Taulí, Sabadell, Barcelona, Spain

<sup>e</sup> Sección de Gestión de Calidad de la Asociación Española de Cirujanos, Servicio de Cirugía, Fundación Hospital de Calahorra, Calahorra, La Rioja, Spain

<sup>f</sup> Hospital General Universitario Los Arcos del Mar Menor, Pozo Aledo, Murcia, Spain

## ARTICLE INFO

### Article history:

Received 18 November 2015

Accepted 27 June 2016

Available online 19 November 2016

### Keywords:

Health information systems

Quality indicators

Surgery department

General surgery

Benchmarking

## ABSTRACT

**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.

<sup>☆</sup> Part of this publication was presented as an oral presentation at the 10th Regional Conference for Quality Care held June 16-17, 2015, under the title: "Pilot study of a set of benchmarking indicators in general surgery".

<sup>☆☆</sup> Please cite this article as: Soria-Aledo V, Angel-Garcia D, Martinez-Nicolas I, Rebas Cladera P, Cabezali Sanchez R, Pereira García LF. Desarrollo y estudio piloto de un conjunto esencial de indicadores para los servicios de cirugía general. Cir Esp. 2016;94:502-510.

\* Corresponding author.

E-mail addresses: ismael.martinez@insp.mx, ismael.m.n@um.es (I. Martinez-Nicolas).

**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.

## Desarrollo y estudio piloto de un conjunto esencial de indicadores para los servicios de cirugía general

### R E S U M E N

#### 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, cuali-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.

**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.

## Introduction

The relationship between Departments of General Surgery (DGS) and quality healthcare is not new. From patient safety<sup>1,2</sup> to improved healthcare processes,<sup>3</sup> there are numerous elements of quality management that departments of surgery have rapidly adapted in order to improve their professional activities. Benchmarking tools and comparative learning have been proposed as driving forces to improve our specialty.<sup>4</sup> However, in spite of indicators that could potentially be useful to this end,<sup>5-7</sup> in surgery there are still few comparative studies between hospital surgical divisions. The only DGS benchmarking articles published are limited to major ambulatory surgery<sup>6</sup> and thoracic surgery.<sup>8</sup>

Currently, there is no balanced set of indicators to compare and facilitate continuous improvements to quality in DGS. Thus, the *Asociación Española de Cirujanos* (AEC-Spanish Association of Surgeons), through their Quality Management Division, has started this initiative. With the ultimate objective of creating a set of indicators for monitoring and

benchmarking DGS at acute care hospitals within the Spanish National Healthcare System (NHS), this article presents the selection, development and results of the pilot study of an initial group of indicators.

## Methods

In this study, 3 successive phases were followed: (1) identification of indicators for quality of care and their prioritization by means of consensus methods; (2) creation and adaptation of datasheets for the selected indicators; and (3) pilot study of the indicators and instruments in cases from the general and digestive surgery department of a public hospital. These stages took place between July and December 2014.

### Identification of the Initial Block and Prioritization of the Indicators

The objectives of this phase were the identification of a group of indicators that: (a) include aspects relative to the

entire healthcare process; (b) represent basic aspects that are able to evaluate the quality of the care provided by a DGS at any acute care hospital within the NHS; (c) have clear scientific evidence or strong agreement among experts; and (d) allow measures for improvement to be implemented, based on their results.

A search was conducted in the database of the Agency for Healthcare Research and Quality (AHRQ) and the databases of Medline, EMBASE and The Cochrane Library from 2000 to 2014, inclusive, using the search terms: “quality indicators” AND “general surgery”; and “quality indicators” AND “digestive surgery”. A RAND-type consensus and prioritization method (modified Delphi)<sup>9</sup> was established, with the participation of a group of 8 surgeons belonging to the Quality Management Division of the AEC and 2 experts in Care Quality from the Healthcare and Social Policy Administration of the Region of Murcia. Three consensus rounds were established (2 by e-mail and one meeting in person), as shown in Fig. 1.

#### Creation and Adaptation of the Indicator Datasheets

A consultation was requested with 2 external experts to organize the indicators in a standardized format (Fig. 2). In this process, the original indicator datasheets were adapted and translated to Spanish (if necessary), while their specifications

were fine-tuned. To this end, the consultants composed several drafts that were reviewed and revised by experts of the workgroup until there were no more comments or modifications.

#### Pilot Study

This phase took place at the Hospital General Universitario Morales Meseguer in Murcia, which is a public hospital with 340 beds and an assigned population of more than 280 000 inhabitants. The indicators were studied separately according to their data source.

#### Indicators Based on Patient Files

Three independent evaluators conducted an external retrospective evaluation. Two independent evaluators (E1 and E2) collected the data simultaneously, while a third evaluator (E3), a surgeon trained only with the materials generated for the pilot study (datasheets and data collection forms), did the same 2 months later.

For each indicator, a sample of 30 cases was contemplated, which was selected by a systematic random sampling from the first semester of 2014. For the identification of cases, surgical discharge lists from the Minimum Basic Data Set (MBDS) were used, using the criteria described in the datasheets for each indicator.<sup>10</sup>

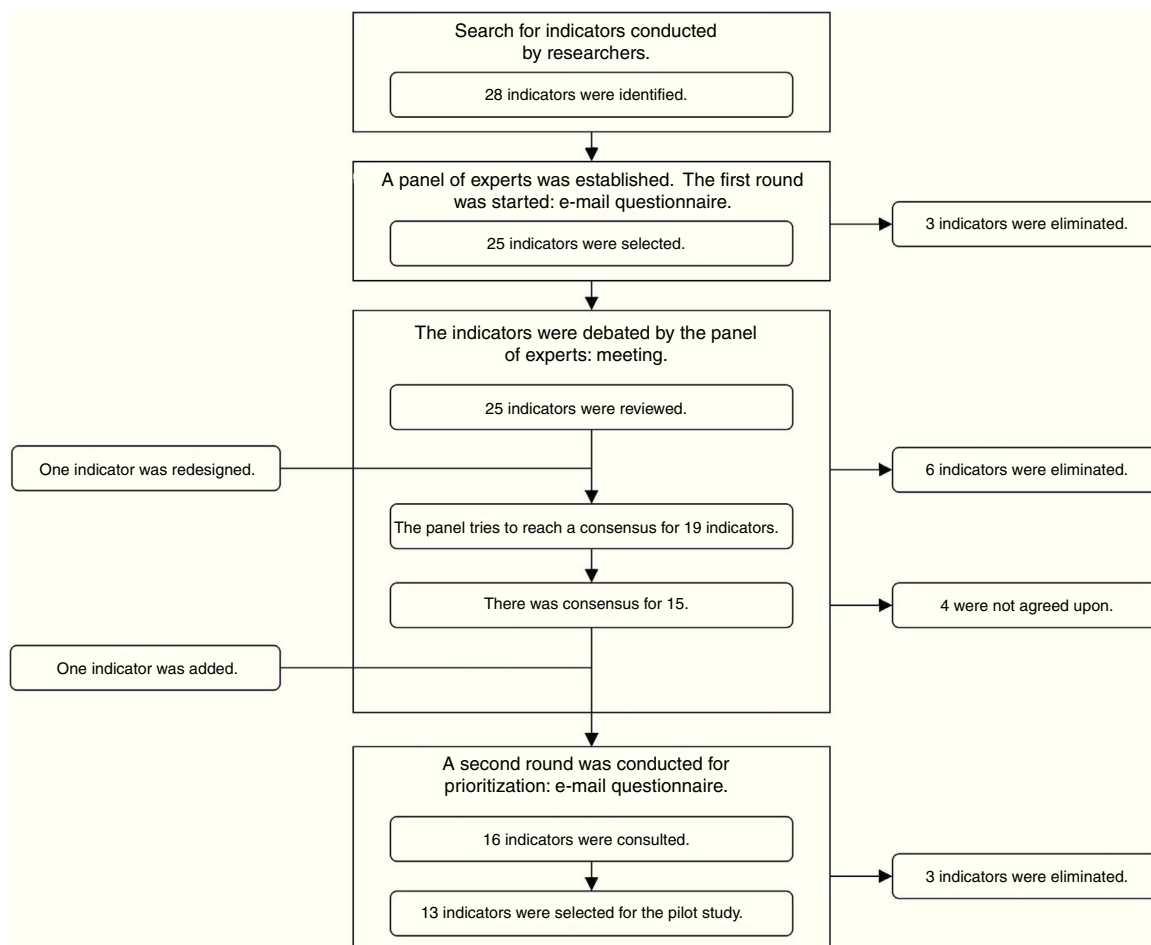


Fig. 1 – Indicator selection and prioritization process.

(Name of indicator)
Set of indicators: (Name of Set)
Subgroup:
Indicator ID #:
Abbreviated name: (Name abbreviated in formula)
Description:
Justification:
Type of measure: (Structure, Process, Result)
Improvement identified as: (Increase/Decrease of result)
Numerator:
Populations included:
Populations excluded:
Data elements: (Data used to identify the numerator)
Denominator:
Populations included:
Populations excluded:
Data elements: (Data used to identify the denominator)
Risk adjustment:
Focus of the data collection: (Evaluation method; data source)
Precision of data:
Suggestions in the analysis of the measurement:
Sampling: (Type of sample; suggested sample size)
Data reported as: (Type of statistical data)
Origin of the indicator:
Selected references:
Notes: (Notes or observations for correct evaluation of the indicator)

**Fig. 2 – Standardized indicator form.**

Three attributes were examined for this type of indicators: feasibility, reliability, and utility to identify opportunities for improvement based on the estimated compliance.

*Feasible Measurement.* We considered feasible those indicators that could be evaluated with the information available at the hospital, reaching the established sample size ( $n=30$ ). Additionally, a qualitative analysis was done with the information collected during the evaluation.

*Interobserver Reliability.* In order to assess the reliability of the indicators, the following steps were followed: (1) analysis of the rate of general agreement ( $P_o$ ) among the 3 evaluators—percentage of cases with the same response—, which should be  $\geq 95\%$  to accept the indicator as reliable; (2) interpretation of the inter-rate reliability with the kappa index ( $\kappa$ ) for 3 evaluators<sup>11</sup>: only those with kappa values  $>0.6$  were considered acceptable, in accordance with the Landis and Koch criteria<sup>12</sup>; and (3) determination of whether the prevalence of the characteristic evaluated was extreme (compliance or noncompliance  $\geq 85\%$ ) as the possible origin of the deviation of  $\kappa$  toward 0, which would not allow the results to be interpreted reliably.<sup>13</sup>

To affront this problem, as well as to observe differences between a measurement made by experienced evaluators and a trained surgeon, the former were analyzed by each of the evaluators. We also we added the calculation of the kappa index adjusted for prevalence-adjusted bias-adjusted kappa (PABAK),<sup>14</sup> whose values were interpreted the same as  $\kappa$ .

Finally, a qualitative analysis was completed with information collected during the evaluation and consultation with evaluators to obtain better knowledge of the existing reliability problems.

*Utility for Identifying Opportunities for Improvement.* We obtained an estimate for the level of compliance of all the indicators and subindicators, as well as their 95% confidence interval, utilizing formulas for non-stratified systematic random sampling. As the indicators were compound, 100% compliance<sup>15</sup> and binomial in nature, the same estimation formulas were applied as in simple indicators and subindicators.

#### *Indicators Based on Minimum Basic Data Sets (MBDS)*

The official hospital records were explored, specifically surgical discharges from 2013. We requested from the Documents Department exported data from the MBDS (both

in hospital care as well as specialized ambulatory care) with variables described in a methodological manual.<sup>10</sup> The definitions of these variables can be consulted on the statistics website of the Spanish Ministry of Healthcare, Social Services, and Equality.<sup>16,17</sup>

With these data, we explored the feasibility and utility of these indicators.

**Feasibility Measurement.** Indicators were considered feasible if they could be calculated with the information available at the hospital, without requiring an external database. The calculation method of each is explained separately.<sup>10</sup>

**Utility of Identifying Opportunities for Improvement.** The prevalence of the indicators was obtained based on their calculation. The calculations were made by an evaluator and later reviewed by another in order to minimize errors.

### Statistical Analysis

All the analyses were carried out with programmed spreadsheets, Stata 13<sup>18</sup> and EpiDat 4.1.<sup>19</sup>

## Results

### Identification of the Initial Set and Prioritization of the Indicators

In the initial search, 28 indicators were identified. After the 3 rounds of prioritization, 16 indicators were ruled out and another was added (Fig. 1). In the end, 13 indicators were proposed for the pilot study (Table 1): 5 indicators from the review of the medical files, and 8 indicators from the electronic database (MBDS).

### Creation and Adaptation of Indicator Datasheets

The 13 indicators were created to meet all the necessary technical specifications based on existing indicators or the available evidence. Four were adapted from the *Patient Safety Indicators of the AHRQ*,<sup>20</sup> the *National Hospital Inpatient Quality Measures* by the *Joint Commission*<sup>21</sup> or the *Key Indicators of the NHS*<sup>22</sup>; 5 were modified from proposals in a previous study of the Spanish Society for Quality Care,<sup>5</sup> and 3 were newly created from different sources. The datasheets, together with the bibliography and the origin of each indicator, are accessible on the web page of the AEC.<sup>10</sup>

### Pilot study

#### Indicators Based on Medical Files

**Feasibility Measurement.** We reviewed a total of 144 patient files until completing 30 study units for each of the 5 indicators. The difficulties found in the evaluation were: 2.2 (*Diagnosis-Treatment Interval in Colorectal Cancer*), numerous exclusions; 3.1 (*Antithrombotic Prophylaxis*), identification of heparin prescription that is difficult to locate in the records during hospitalization days in ICU/Recovery or the day of discharge; 3.2 (*Antibiotic Prophylaxis*), several data sources: time

of administration, time of start of surgery, antibiotic, and dose in different records.

**Interobserver Reliability.** The results of the quantitative analysis for reliability are shown in Table 2. The indicators for the interval between diagnosis and treatment, both in colon cancer as well as breast cancer, have been shown to be reliable with a  $\kappa > 0.6$ . The indicator for the informed consent document (ICD) was evaluated with PABAK (due to extreme prevalences), and was also satisfactory in all its components.

As for the prophylaxis indicators,  $\kappa$  lower than 0.6 were obtained (no extreme prevalence), so it cannot be concluded that these indicators are reliable. These indicators are affected by the sum of small disagreements of all their subindicators.

The qualitative analysis showed: (a) the difficulty to adequately count the days of prescribed antithrombotic agents, as well as the interpretation of the type of drug and proper dosage on all the days of hospitalization; and (b) the records may not be precise enough to exactly determine the actual antibiotic administration time before surgery.

**Estimation of Compliance.** Table 3 shows the estimation for these indicators. The compound ICD indicator obtained very low results. It was found that the ICD was present and corresponded with the surgery (93.33% and 100%, respectively), but the subindicators regarding physician data (3.33%) caused this low compliance.

#### Indicators Based on Minimum Basic Data Sets

**Feasibility of the Measurement.** All the indicators could be calculated based on the MBDS. However, the indicator 4.7 (*Reoperation after scheduled surgery*) needed to be adapted from "reoperation" to "readmission" as it was not possible to correctly identify the reoperations during one same health-care episode. Meanwhile, indicator 4.2 (*Readmission in less than 7 days after Major Ambulatory Surgery*) required additional information from the MBDS for hospital care (MBDS-HC) and specialized ambulatory care (MBDS-SAC).

**Prevalence of Surgical Events.** The results obtained are shown in Table 4 as proportions.

## Discussion

This article explains the selection and pilot study of an initial set of DGS indicators for a project that will later be developed. A set of 13 indicators were selected, which were able to be measured with available data sources. Only one MBDS indicator (*Reoperation within 30 days after scheduled surgery*) had serious limitations for measurement and may be considered unfeasible. Out of the 5 indicators from patient files, 3 can be considered reliable, while the 2 indicators for antibiotic and antithrombotic prophylaxis were redesigned with the results of this study.

In recent years, interest has grown in measuring and improving quality in DGS, and one tool used to this end is benchmarking.<sup>23</sup> Previous studies have demonstrated its utility in the field of surgery.<sup>4,24,25</sup> Nonetheless, in Spain there still is not a good system of indicators in place that could be

**Table 1 – Indicators Proposed for General Surgery Departments.**

Proposed indicators	Numerator	Denominator	Data source
1.1 Correct completion of the informed consent document	Patients who were treated surgically by the general surgery department and correctly completed the informed consent document	Patients who were treated surgically by the general surgery department	Patient files
2.1 Preoperative hospitalization less than 1 day	Patients who had surgery the same day or the day following admission (preoperative stay $\leq 1$ day)	Patients hospitalized for scheduled surgery	MBDS
2.2 Interval between diagnosis and therapy in colorectal cancer	Patients with colorectal cancer whose interval between diagnostic confirmation and surgical treatment was $\leq 28$ calendar days	Patients diagnosed with colorectal cancer who received surgical treatment	Patient files
2.3 Interval between diagnosis and therapy in breast cancer	Patients with breast cancer whose interval between diagnostic confirmation and initial surgical treatment was $\leq 4$ weeks or 28 calendar days	Patients diagnosed with breast cancer whose initial treatment was surgical	Patient files
3.1 Thromboembolic prophylaxis in colorectal cancer	Patients treated for colorectal carcinoma with bowel and/or rectal resection, with correct thromboembolic prophylaxis <sup>a</sup>	Patients treated surgically for colorectal carcinoma with bowel and/or rectal resection	Patient files
3.2 Antibiotic prophylaxis in colorectal cancer	Patients treated for colorectal carcinoma who undergo bowel and/or rectal resection with correct antibiotic prophylaxis <sup>a</sup>	Patients treated surgically for colorectal carcinoma with bowel and/or rectal resection	Patient files
4.1 Readmission within 30 days after surgery	Patients with unscheduled admission within an interval of 30 days after hospital discharge	Patients treated with either scheduled or urgent surgery	MBDS
4.2 Readmission within 7 days after major ambulatory surgery	Patients admitted within 7 days after MAS	Patients with scheduled surgical interventions in the general surgery DRG/MAS setting	MBDS
4.3 Postoperative hemorrhage or hematoma (reoperation)	Cases of hemorrhage or hematoma that require a surgical procedure to control the hemorrhage	Surgical discharges with codes for surgical procedures; DRG surgical discharges with date of intervention	MBDS
4.4 Abdominal wall dehiscence after surgery	Discharges of patients $>14$ years old with code ICD-9 for closure of a disruption of the abdominal wall	Discharges from abdominal surgery in patients $>14$ years of age	MBDS
4.5 Infection in clean surgery	Discharges with surgical site infection in any secondary diagnostic field	Scheduled discharges after clean surgery (breast surgery, hernia surgery, thyroid, and parathyroid surgery)	MBDS
4.6 Infection in colorectal surgery	Discharges with infection of the surgical site in any secondary diagnostic field	Scheduled discharges with colorectal surgery	MBDS
4.7 Reoperation after scheduled surgery	Patients who underwent unscheduled surgery during the same hospitalization episode within 30 days of previous surgery	Total number of patients who underwent programmed surgery	MBDS

ICD, international classification of diseases; MAS, major ambulatory surgery; MBDS, Minimum Basic Data Set; DRG, diagnosis-related group.

<sup>a</sup> The definition of “corrective” prophylaxis is found on the datasheet for each indicator.<sup>10</sup>

used. A review of the existing literature and consensus methods that have been widely accepted and used within the scientific community<sup>26-28</sup> has allowed us to select highly relevant indicators for DGS.

This study continues along the same lines of previous studies and research. Thus, our set of indicators includes, among others, 5 indicators from a prioritization document by the Spanish Society for Quality Healthcare<sup>5</sup> that have not been used in pilot studies or been evaluated to date. Given the undeniable need for empirically studying indicators before their utilization,<sup>29-31</sup> we have completed a pilot study following the methodology applied in previous projects<sup>32</sup> that also incorporates a greater analysis of interobserver reliability.

An essential aspect for comparative indicators is the reliability and consistency of their measurement. We have therefore presented kappa and PABAK results, as recommended by other authors when there are extreme prevalences and based on the argument that only one concordance index is insufficient.<sup>14,33,34</sup> The compound indicators for prophylaxis have been the most conflictive, as they are conditioned by their subindicators and the accumulation of their errors. One study limitation could be the sample size used. There is not sufficient consensus on this subject, although some authors recommend no less than 30 observations.<sup>35</sup>

On the other hand, the compliance results show opportunities for improvement in the pilot hospital, which supports the usefulness of these indicators. Thus, the inadequate

**Table 2 – General Agreement, Kappa Index, PABAK, and Prevalence of Each Indicator for the 3 Evaluators or by Pairs.**

Indicator	3 evaluators			Eva1-Eva2			Eva1-Eva3			Eva2-Eva3		
	P <sub>o</sub>	κ (IC 95%) <sup>a</sup>	P% <sup>b</sup>	PABAK	κ	P% <sup>b</sup>	PABAK	κ	P% <sup>b</sup>	PABAK	κ	P% <sup>b</sup>
1.1. ICD	93%	0.58	94%	<b>0.87</b>	0.47	93%	<b>0.87</b>	0.47	93%	1	1	97%
Present in the patient file	<b>97%</b>	0.74	94%	0.93	0.65	95%	0.93	0.65	95%	1	1	97%
ICD corresponding with the intervention	<b>100%</b>	–	100%	1	–	100%	1	–	100%	1	–	100%
3 criteria:	93%	0.57	94%	<b>0.86</b>	0.47	93%	<b>0.86</b>	0.47	93%	1	1	96%
• Signature of doctor and medical board n	93%	0.57	94%	<b>0.86</b>	0.47	93%	<b>0.86</b>	0.47	93%	1	1	96%
• Patient/proxy signature	89%	–0.04	96%	<b>0.93</b>	0	98%	<b>0.86</b>	0	96%	0.79	–0.05	95%
• Place, date	<b>100%</b>	–	100%	1	–	100%	1	–	100%	1	–	100%
2.2. Interval between diagnosis and treatment, colon	83%	<b>0.77 (0.57–0.96)</b>	60%	1	1	57%	0.67	0.65	62%	0.67	0.65	62%
2.3. Interval between diagnosis and treatment, breast	90%	<b>0.63</b>	90%	0.93	0.78	92%	0.8	0.52	88%	0.87	0.63	90%
3.1. Antithrombotic prophylaxis	73%	0.56 (0.29–0.82)	72%	<b>0.8</b>	0.78	65%	<b>0.67</b>	0.53	78%	0.47	0.37	73%
Daily prescription, hospitalization	77%	0.14	90%	<b>0.67</b>	0.36	85%	<b>0.8</b>	0	95%	0.6	0	90%
Adequate medication	<b>100%</b>	–	100%	1	–	100%	1	–	100%	1	–	100%
Adequate dosage	90%	0.53	92%	<b>0.87</b>	0.63	90%	<b>0.93</b>	0.65	95%	<b>0.8</b>	0.37	92%
Prescription at discharge +30 days	93%	<b>0.81</b>	86%	0.93	0.87	85%	0.87	0.71	87%	0.93	0.84	88%
3.2. Antibiotic prophylaxis	70%	0.56 (0.31–0.81)	64%	<b>0.75</b>	0.72	60%	0.4	0.36	65%	<b>0.67</b>	0.62	68%
Adequate medication	93%	0.48	96%	<b>1</b>	1	93%	<b>0.87</b>	0	97%	<b>0.87</b>	0	97%
Adequate dosage	<b>96%</b>	–0.04	99%	0.93	0	98%	0.93	0	98%	1	–	100%
Previous administration (1 h)	77%	<b>0.64 (0.40–0.89)</b>	68%	0.8	0.78	65%	0.53	0.47	68%	0.73	0.68	70%

Eva1, evaluator 1; Eva2, evaluator 2; Eva 3, evaluator 3; κ, kappa index; P<sub>o</sub>, rate of general agreement or observed agreement; P%, prevalence; in bold, values presenting good reliability.

<sup>a</sup> Kappa and confidence interval jackknife at 95% for 3 evaluators. The confidence interval is only shown if there is no extreme prevalence.

<sup>b</sup> Prevalence of the most frequent evaluated characteristic (either compliance or noncompliance).

**Table 3 – Estimated Completion of the Indicators and Subindicators From the Review of Patient Files (1st Semester 2014).**

Indicator	Estimation <sup>a</sup>	95%CI
1.1 ICD	3	±5
Present	93	±7
ICD correct	100	±0
3 criteria	3	±5
Doctor	3	±5
Patient or proxy	93	±7
Place and date	100	±0
2.2 D-T interval, colorectal cancer	57	±15
2.3 D-T interval, breast cancer	97	±9
3.1 Thromboembolic prophylaxis	63	±10
Prescribed the entire hospitalization	77	±9
Adequate medication	100	±0
Adequate dosage	87	±7
Prescription at discharge	83	±8
3.2 Antibiotic prophylaxis	57	±11
Adequate medication	93	±5
Adequate dosage	90	±6
Previous administration (1 h)	63	±10

ICD, informed consent document; D-T, diagnosis-treatment; CI, confidence interval.

<sup>a</sup> Estimation of the percentage of compliance with confidence interval for proportions with non-stratified systematic randomized sampling.

completion of the ICD shines a new light on studies done in this setting,<sup>36–38</sup> as these focus on the formal aspect and not on the use of these documents. Furthermore, the low number of colorectal cancer cases that comply with the recommended time interval between diagnosis and treatment (57%; 95% CI ±15%) is similar to the results from the NHS Cancer Strategy from 2009<sup>39</sup> (54.7%; 95% CI: 51.9%–57.4%), while in breast cancer these levels are much better (97%, 95% CI ±9% vs 43.6%, 95% CI: 41%–46%). Nevertheless, the results for the MBDS indicators require proper comparison with a standard of reference, either external or from the hospital itself, in order to determine areas for improvement.

For this comparison to be valid in result indicators, it is necessary to analyze the precision of the indicators and adjust them according to the case-mix of the population treated.<sup>40,41</sup> Moreover, we have shown that it is feasible to construct and measure MBDS indicators, which enables us to continue researching their precision as well as any necessary adjustments for their comparison between hospitals.

In conclusion, the set of 13 indicators aimed at benchmarking DGS that are presented in this article have been demonstrated to be feasible, with the exception of one. Two of the indicators from patient files have been revised and re-evaluated to ensure validity. More research is required for the adjustment of risk for the result indicators as well as in the calculation and automatization of the MBDS indicators. Five indicators can currently be used to search for opportunities for

**Table 4 – Prevalence in Indicators of MBDS for Discharges From the Surgery Department at the Hospital General Universitario Morales Meseguer (2013).**

Indicator	Events	Denominator	Prevalence
2.1 Preoperative hospital stay $\leq$ 1 day	993	1000	99.3%
4.1 Readmission <30 days after surgical intervention	95	1481	6.4%
4.2 Readmission <7 days after MAS	5	365	1.4%
4.3 Postoperative hemorrhage/hematoma (reoperation)	40	1000	4%
4.4 Abdominal wall dehiscence	8	454	1.8%
4.5 Infection after clean surgery	2	414	0.5%
4.6 Infection after colorectal surgery	11	135	8.1%
4.7 Reoperation after scheduled surgery	4	927	0.4%

improvements and to make comparisons between hospitals without the need for any adjustments.

### Funding

Financed by the Spanish Association of Surgeons: "AEC Project to define indicators".

### Authorship

Victor Soria-Aledo contributed to the initial concept and design of the study, composition of the article and final approval before publication. Daniel Angel-Garcia contributed to the collection, analysis and interpretation of the data, composition and critical review of the article and final approval before publication. Ismael Martinez-Nicolas contributed to the study design and field work, collection, analysis and interpretation of the data, composition of the article and final approval before publication. Pere Rebasá Cladera contributed to the initial concept and design of the study, participation and coordination of the group of experts, critical review with important intellectual offerings and approval of the final version for publication. Roger Cabezali Sanchez contributed to the initial concept, participation of the group of experts, composition, critical review and approval of the final version for publication. Luis Francisco Pereira García contributed to the data collection and analysis, composition, critical review and approval of the final version for publication.

### Conflict of Interests

The authors declare having no occupational, research, economic or ethical conflicts of interests regarding this article. The conclusions of this document are absolutely independent, and no external agents have influenced them.

### Acknowledgements

The authors would like to acknowledge the following: the EMCA Program of the Healthcare Administration in the Region of Murcia, Spain; Pedro Parra Hidalgo and José Eduardo Calle Urra, experts in quality care in the Region of Murcia, for their assessment and collaboration in the indicator development

process; the Spanish Association of Surgeons (AEC) for their involvement, support and funding; and the experts who participated in the indicator selection and prioritization phase.

### REFERENCES

- De Vries EN, Prins HA, Crolla RMPH, den Outer AJ, van Anel G, van Helden SH, et al. Effect of a comprehensive surgical safety system on patient outcomes. *N Engl J Med*. 2010;363:1928–37.
- 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.
- Cevasco M, Ashley SW. Quality measurement and improvement in general surgery. *Perm J*. 2011;15:48–53.
- 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.
- Sociedad Española de Calidad Asistencial. Indicadores de calidad para hospitales del Sistema Nacional de Salud [Internet]; 2012;111. Available in: <http://www.calidadasistencial.es/> [accessed 14.02.16]
- 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.
- 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.
- 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.
- Fink A, Kosecoff J, Chassin M, Brook RH. Consensus methods: characteristics and guidelines for use. *Am J Public Health*. 1984;74:979–83.
- 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. Available in:



- <http://www.aecirujanos.es/wpaec/actualizacion-disponible-set-indicadores-calidad-los-servicios-cirugia-general/> [accessed 05.01.16]
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. Available in: <http://pestadistico.inteligenciadegestion.msssi.es> [accessed 26.03.16]
  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. Available in: <http://pestadistico.inteligenciadegestion.msssi.es/> [accessed 26.03.16]
  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). Epidat: programa para análisis epidemiológico de datos. Colombia: Universidad CES; 2014. Available in: <http://dxsp.sergas.es/> [accessed 14.07.15]
  20. Agency for Health Care Research and Quality (AHRQ). Patient safety indicators technical specifications updates—version 5.0, March 2015 [Internet]. AHRQ quality indicators; 2015. Available in: [http://www.qualityindicators.ahrq.gov/Modules/PSI\\_TechSpec.aspx](http://www.qualityindicators.ahrq.gov/Modules/PSI_TechSpec.aspx) [accessed 26.03.16]
  21. The Joint Commission. Specifications manual for national hospital inpatient quality measures [Internet]. The joint commission: performance measurement; 2015. Available in: [http://www.jointcommission.org/specifications\\_manual\\_for\\_national\\_hospital\\_inpatient\\_quality\\_measures.aspx](http://www.jointcommission.org/specifications_manual_for_national_hospital_inpatient_quality_measures.aspx) [accessed 26.03.16]
  22. Ministerio de Sanidad Servicios Sociales e Igualdad (MSSSI). Indicadores Clave Sistema Nacional de Salud. Versión 1.0 [Internet]; 2016. Available in: <http://inclasns.msssi.es/> [accessed 26.03.16]
  23. Mace SE. Continuous quality improvement for the clinical decision unit. *J Health Care 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, Alverti 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, Ángel-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.