

La evaluación local de la calidad mediante instrumentos de medida genéricos

Vahé A. Kazandjian^a / Paula Vallejo^b

^aCenter for Performance Sciences, Elkrige, Maryland. Adjunct Professor, Bloomberg School of Public Health, The Johns Hopkins University, Baltimore, Maryland, Estados Unidos.

^bTesser-Land, Madrid, España.

Correspondencia: Vahe A. Kazandjian, Center for Performance Sciences, 6820 Deerpath Road, Elkrige, Maryland 21075, USA. Correo electrónico: vkazandjian@cpsciences.com

Recibido: 5 de junio de 2003.
Aceptado: 2 de febrero de 2004.

(Local evaluation of quality using generic measurement tools)

Resumen

La medición de calidad y la responsabilidad son dos de los conceptos más utilizados en los sistemas sanitarios, aunque los instrumentos para poner en práctica estos conceptos varían. El objetivo de este artículo es analizar las herramientas (p. ej., los indicadores) y las expectativas de los usuarios de estas herramientas (p. ej., la calidad frente al rendimiento) al desempeñar el imperativo de responsabilidad en el sistema sanitario. Concretamente, presentamos la evolución del Sistema Sanitario Español en esta materia para analizar cómo se han ido constituyendo paso a paso los principios del modelo nacional de responsabilidad. Por último, se presenta un sistema internacional, el Maryland Quality Indicator Project[®], como marco de referencia para la evaluación del rendimiento.

Palabras clave: Indicadores. Calidad. Responsabilidad. Sistemas de medición.

Abstract

Performance measurement and accountability are the most frequently encountered concepts across healthcare systems, although the tools used to put these concepts into practice vary. The aim of the present article was to discuss the tools (e.g., indicators) and user expectations (e.g., quality versus performance) of these tools in fulfilling healthcare's accountability mandate. In particular, the developments within the Spanish health system are discussed to illustrate how Spain has addressed, in a stepwise manner, the determinants of a national accountability model. Finally, a multinational project, The Maryland Quality Indicator Project[®], is presented as a popular framework for measuring performance.

Key words: Indicators. Quality. Responsibility. Measurement system.

Introducción

El interés en la medida del rendimiento en hospitales y otras organizaciones sanitarias comenzó a principios de los años ochenta, fundamentalmente en Estados Unidos^{1,2} y Australia³. Los motivos que impulsaron dichas iniciativas fueron la acreditación de hospitales y la confianza en que las medidas de rendimiento, que en un principio fueron diseñadas como cocientes basados en indicadores, servirían como instrumento para que los organismos acreditadores decidieran qué organizaciones estaban actuando «correctamente». Al principio, los indicadores propuestos por la Joint Commission on Accreditation of Health Care Organizations (JCAHO) en Estados Unidos y por el Council for Quality Care en Australia estaban basados en medidas de estructura del hospital, como la frecuencia de comisiones de temas relacionados con el rendimiento, los métodos de comuni-

cación interna y las comisiones de gestión para la revisión del desempeño. Posteriormente, estas medidas de estructura dieron paso a sistemas más «tradicionales», centrados sobre todo en el aspecto clínico del tratamiento. Algunas medidas ampliamente evaluadas en sistemas sanitarios de varios continentes fueron los cocientes de infección hospitalaria, de mortalidad, de cesárea y de readmisiones.

En la actualidad, los sistemas de medición de resultados van más allá de la pretensión inicial de usar indicadores para la acreditación hospitalaria, cuyo objetivo principal era demostrar la responsabilidad ante grupos de interés muy diversos (pacientes, comunidades, gobiernos, empresas aseguradoras y de negocios). Esta evolución puede observarse ahora en Europa, donde las iniciativas llevadas a cabo en el Reino Unido⁴, Alemania⁵, Holanda^{6,7} y Dinamarca⁸ han servido de punto de partida para un sistema europeo de responsabilidad y mejora de la calidad. Al hablar de «res-

ponsabilidad» en el ámbito sanitario (del término anglosajón *accountability*) nos referimos al deber que tiene el proveedor de los servicios de responder sobre su actuación ante el receptor de dichos servicios. En este artículo proponemos la necesidad de medir dicha responsabilidad para que el receptor de los servicios sea capaz de poner un valor a la actuación (o calidad) del servicio.

El objetivo de este artículo es analizar las dificultades principales relacionadas con la medición del desempeño y de la responsabilidad, centrándonos en las actividades que se están llevando a cabo en el sistema sanitario español. Finalmente, se presenta el programa internacional Maryland Quality Indicator Project® como modelo para ilustrar las dificultades y los éxitos obtenidos en la aplicación de una metodología homogénea a gran número de sistemas sanitarios de distintas características.

Aspectos metodológicos previos

El primer concepto que es necesario tratar es la diferencia entre rendimiento y calidad. Los sistemas de evaluación del rendimiento han prometido frecuentemente más de lo que su propio nombre implica. De hecho, mientras «actuación» o «rendimiento» son los términos que definen estas iniciativas, la mayoría de los usuarios potenciales de la información han transformado el término «actuación» en «calidad». En muchas ocasiones, incluso hemos traducido el término anglosajón *performance* (actuación, rendimiento) directamente por calidad. Aunque este artículo considera actuación o rendimiento como la calidad menos el valor añadido por el evaluador, entendemos que su traducción al castellano no puede ser literal.

Las expectativas puestas en los sistemas de evaluación de la actuación, por tanto, no han sido realistas, ya que lo que en realidad identifican son las áreas clave de la actuación de una organización que, a través de un análisis cuidadoso, sistemático y prospectivo, pueden señalar aspectos que requerirían una mayor comprensión y, si fuese necesario, una evaluación de su adecuación.

Se ha formulado que la medida del rendimiento sirve como base para la evaluación de la responsabilidad mientras la medida se mantiene a nivel estadístico, sin un juicio de valor sobre la bondad de lo que se está midiendo⁹. Por ejemplo, un cociente de cesáreas del 25% comparado con uno del 19% no implica directamente que el más alto es inapropiado, si tenemos en cuenta que hay la posibilidad de que los hospitales con un cociente de cesáreas bajo pueden estar potenciando innecesariamente el parto vaginal a mujeres primerizas

o el parto vaginal tras una cesárea a mujeres con cesáreas previas. Para comprender el cociente de cesáreas hay que ir más allá del mero análisis del coste potencial de cada tipo de parto, asociándolo a los resultados obtenidos en la salud tanto de la madre como del niño. Una explicación similar se puede aplicar a la mayoría de los indicadores, incluso a razones de mortalidad, tiempos de espera, retornos a la sala de operaciones o reingresos hospitalarios. Ninguno de estos indicadores implica *a priori* una buena o mala calidad.

Por tanto, la distinción entre actuación y calidad es fundamental, si tenemos en cuenta que actualmente se considera que la calidad es el resultado de la interpretación de los indicadores de actuación, en lugar de los resultados directos de magnitud que reflejan los cocientes¹⁰. Así, mientras que la medida de la actuación se puede realizar de forma genérica a través de indicadores universales, la transformación de éstas en un sistema de valores es más fácil de obtener localmente, según la propia filosofía del lugar acerca de la salud, la enfermedad, el acceso al tratamiento sanitario y la disponibilidad de recursos.

Otro concepto que es importante clarificar es la diferencia entre cocientes y cálculos. Los métodos de medición de actuación se pueden clasificar en 2 categorías generales: los métodos que tienen en cuenta únicamente los datos del numerador y los métodos que están basados en cocientes o proporciones, utilizando datos del numerador y del denominador conjuntamente. La distinción entre estas 2 categorías es necesaria, ya que hay investigaciones en la bibliografía sanitaria que demuestran que las tendencias en la práctica clínica varían de forma más significativa cuando los datos que se muestran a los profesionales sanitarios de forma individual, o a las organizaciones sanitarias en conjunto, son de tipo comparativo. Esta evidencia es el resultado, principalmente, de una serie de proyectos de análisis de variaciones de tendencias de la práctica clínica que se agrupan bajo el nombre de *small variations analysis* (SAVA)¹¹. Estos estudios han demostrado que es necesario indicar a los proveedores dónde se encuentran en relación con otros (especialmente con sus pares), sus procesos y sus resultados en la gestión de actividades específicas del hospital o incluso en la gestión de una enfermedad específica, para conseguir que cambien sus ideas iniciales sobre las tendencias en la práctica clínica y su estilo de práctica, y adopten nuevos métodos y procesos. Aunque la terminología de la medicina basada en la evidencia no era popular durante los primeros años del análisis de variaciones, las conclusiones de esos estudios han servido de germen al proyecto actual de establecer prácticas basadas en la evidencia para eliminar la subjetividad y la opinión personal. En algunos aspectos, mostrar cocientes de indicadores y sus tendencias longitudinales consistentes

supone una estrategia similar a reforzar las prácticas que tienen una relación consistente entre procesos y resultados. En un intento de establecer paralelismos con la medicina basada en la evidencia, podemos afirmar que los análisis de variación que usan el perfil de cocientes de la organización como unidad en realidad logran un rendimiento organizacional basado en la evidencia, donde ésta es tangencial o incluso colateral en forma de *ratio*, o tendencia de resultados entre organizaciones, que muestra la diferencia entre los cocientes.

Hay todavía otro concepto pertinente para comprender lo que los sistemas de medida de actuación pueden lograr y lo que nos prometen implícitamente: la distinción entre indicador e índice. Los indicadores, por definición, tienen que indicar. Aunque decir esto parezca una trivialidad, a menudo se malinterpreta las capacidades de los indicadores. Como medida, un indicador es capaz de señalar las áreas de actuación que parecen necesitar atención, ya que muestran tendencias que podrían indicar un nivel de actuación inaceptable. Teniendo en cuenta que «indicar» no es lo mismo que «interpretar», el mejor uso —de hecho el único uso— de los indicadores está asociado a su capacidad de centrar el foco de análisis en la actuación de una organización o un individuo. Por tanto, los indicadores son útiles cuando se tiene objetivos bien definidos para las preguntas. Una metáfora sobre los indicadores sería la de un perro de caza que se lleva al campo para «indicar» al cazador dónde está escondido el pájaro¹². El mensaje de esta metáfora es que un buen perro de caza podrá ayudar al cazador a identificar dónde está el objetivo de la caza (en este caso el pájaro), pero el perro nunca será capaz de cazar al pájaro. En último término, el éxito de la cacería dependerá de la habilidad del cazador para interpretar de forma correcta el mensaje de su perro, y de su puntería. Los indicadores son precisamente eso: perros de caza disfrazados. Cuando el usuario de los datos proporcionados por los indicadores está entrenado para interpretar correctamente el mensaje, los sistemas de medición de la actuación y los usuarios de dichos sistemas identificarán, de forma conjunta, las áreas cuyos resultados de rendimiento sugieren una atención posterior.

El objetivo de un índice es muy distinto. Los índices tienen 2 características: *a)* están compuestos por múltiples indicadores, y *b)* llevan implícito un juicio de valor. Los índices se han usado tradicionalmente en economía y otras ciencias sociales¹³⁻¹⁶. Cuando se analizan los indicadores financieros, éstos señalan directamente al usuario si la actuación financiera evaluada ha sido buena o mala. Por ejemplo, los índices de productividad, de velocidad o de retorno de la inversión (ROI) llevan implícito un juicio de valor. Cuando se tiene en cuenta varios índices, será necesaria la habilidad y el conocimiento del intérprete para identificar el mensaje subyacente al conjunto de índices, pero cada índice de

forma individual tendrá en sí mismo la función (si es un índice válido) de describir inmediatamente la mayor o menor adecuación de la actuación.

La distinción entre índice e indicador es importante, principalmente porque la magnitud y la dirección que nos muestra el índice están más cercanas al concepto de calidad de la actuación. En el caso de los indicadores, sin embargo, no existe un juicio de valor (tal como ha quedado reflejado en el ejemplo del cociente de cesáreas expuesto previamente) asociado con la magnitud o la dirección que muestran. Esta explicación, por tanto, ilustra por qué los sistemas de medida de actuación sanitaria basados en indicadores no pueden medir el grado de bondad de la actuación, y cualquier expectativa en este sentido sólo servirá para frustrar y confundir a quienes revisan los perfiles de los indicadores.

La responsabilidad del sistema sanitario español con los ciudadanos

El sistema sanitario español lleva más de 2 décadas trabajando en los procesos de mejora de la calidad y la responsabilidad. La evolución de estos conceptos queda patente en los pasos seguidos desde el primer plan de mejora de la calidad del país, establecido en el Hospital de la Santa Creu i Sant Pau de Barcelona en 1981, en los principios de la Ley General de Sanidad y en la iniciativa del Instituto Nacional de Salud de introducir la calidad como un requisito de los contratos de gestión firmados anualmente con los hospitales. De hecho, la Ley General de Sanidad de 1986 puede considerarse un hito decisivo al articular estos conceptos en una ley general sanitaria, estableciendo la necesidad de evaluación de los servicios sanitarios y la promoción de iniciativas de mejora de la calidad.

Entre los impactos de la Ley General de Sanidad está la reestructuración de los servicios sanitarios y algunas iniciativas de evaluación de servicios en varias zonas de la geografía española, realizadas durante el proceso progresivo de transferencias sanitarias. Esta nueva organización del sistema sanitario tiene varias implicaciones para la evaluación y la mejora de calidad. Algunos de los aspectos positivos¹⁷ están relacionados con la mayor inclusión de las opiniones del ciudadano sobre el sistema sanitario. Entre los aspectos potencialmente negativos de las transferencias, está la posibilidad de crear un modelo sanitario que, si no se diseña cuidadosamente, puede dificultar la cooperación entre los distintos sistemas de gestión y la compatibilidad de sistemas de información.

Tras el traspaso de responsabilidades a las comunidades autónomas, se define la primera ley específica de calidad, la Ley de Cohesión y Calidad del Sistema Nacional de Salud, publicada en 2003.

La iniciativa precursora de la mejora de calidad en España fue la publicación del «Plan de Humanización de Servicios» por el Instituto Nacional de Salud (Insalud), en 1984. En 1986, el Insalud publicó el Plan de Garantía de Calidad Total, estableciendo objetivos para el sistema sanitario nacional y para las instituciones sanitarias. El Plan de Calidad está estratificado en objetivos generales y específicos. Cada centro sanitario debe acordar estos objetivos estableciendo la meta específica para la organización, que quedará formalmente fijada en el Contrato de Gestión, firmado cada año por los gerentes de las instituciones. Cada objetivo tiene un indicador, una definición y una fórmula que están previstos en el plan. Los datos necesarios para los indicadores se recogen de 2 maneras: unos parten del CMBD (conjunto mínimo de datos básicos), y otros tienen que ser específicamente recogidos por los hospitales.

Aparte de las iniciativas nacionales, cada comunidad autónoma ha desarrollado o adoptado sus propias leyes de ordenación sanitaria, modelos de calidad y, en algunos casos, sus propios modelos de acreditación. Varias comunidades autónomas están adoptando el modelo de la European Foundation for Quality Management (EFQM) sobre la gestión y la autoevaluación de los servicios sanitarios (p. ej., el País Vasco y la Comunidad de Madrid). Otras regiones han sintetizado varios modelos disponibles en el sector. Por ejemplo, Cataluña ha definido un modelo de acreditación multidimensional aplicable a todos los servicios sanitarios, basado en el modelo EFQM, y combina las normas de la Joint Commission on Accreditation of Healthcare Organizations (JCAHO) y el protocolo ISO. A su vez, Andalucía creó la Agencia de Calidad Sanitaria, que será responsable de la acreditación de los centros sanitarios andaluces siguiendo su propio modelo de acreditación, compuesto por más de 400 estándares que incluyen varios niveles de atención (ciudadano, organización de la actividad, profesionales, estructura y resultados), y está basado en la definición y evaluación de procesos críticos¹⁸. Finalmente, en Galicia se publicó un libro de normas para la acreditación hospitalaria estratificada por servicios.

El desarrollo y la investigación en los distintos modelos de acreditación y certificación de calidad (especialmente EFQM, ISO y JCAHO) han estado unidos a su aplicación en distintas regiones y sectores. En muchas regiones de España, el EFQM se ha establecido como modelo oficial que guía las actividades de mejora de calidad en el sector sanitario, a pesar de la dificultad de su terminología para los profesionales sanitarios¹⁹ y la dificultad de comprensión de criterios y subcriterios²⁰. Estos motivos potenciaron el inicio de investigaciones para adaptar el modelo al campo sanitario.

Un grupo de expertos en gestión de calidad de varias instituciones sanitarias españolas realizó el proyecto

de investigación «Guía para la Autoevaluación de Centros Sanitarios» utilizando como referencia el Modelo EFQM²¹, financiado por el Fondo de Investigación Sanitaria del Instituto de Salud Carlos III. El resultado del proyecto es un documento basado en el consenso que explica el EFQM con un lenguaje adaptado a los profesionales sanitarios. Para los criterios de los resultados se dan algunos ejemplos de los conceptos que deben medirse como indicadores, pero no se definen éstos ni el método de recogida de datos.

Teniendo en cuenta que el EFQM es un modelo no prescriptivo y, por tanto, no especifica la metodología que debe emplearse para la evaluación de los resultados, la comparación entre las distintas instituciones no ha sido fácil. Por ejemplo, el gobierno vasco está llevando a cabo varios proyectos para obtener los datos comparativos de las instituciones sanitarias. Uno de ellos sería la publicación en 2001 de un manual de evaluación y mejora de la satisfacción de los trabajadores de las organizaciones de servicios²².

El proyecto de identificación de un conjunto de indicadores clave de calidad en el ámbito hospitalario según el EFQM²³, publicado en 2001, tenía como objetivo definir los indicadores que midiesen los criterios y subcriterios del EFQM en el sector sanitario. Los resultados de la investigación son limitados ya que, aunque se definen 106 indicadores, no se especifica la fórmula que debe utilizarse ni el método de recogida de datos. Un proyecto similar intentó crear un cuadro de mandos hospitalario basado en el Modelo de Kaplan y Norton y adaptado al modelo EFQM²⁴. La investigación se hizo mediante el método de Delphi y el resultado es un total de 66 indicadores divididos en 6 categorías: los originales de Kaplan y Norton, práctica clínica y actuación social de la organización.

Potenciado por la normativa legal de algunas comunidades autónomas y los resultados de las investigaciones realizadas, el EFQM está siendo ampliamente aplicado en atención primaria²⁵, los hospitales²⁶⁻²⁹ y algunos servicios regionales de salud³⁰ como método de autoevaluación o de evaluación externa. En España, el Club Gestión de Calidades es la organización que coordina las iniciativas de EFQM y la evaluación externa con este modelo.

El protocolo ISO surgió en 1987 como un conjunto internacional de normas de calidad con el objetivo de establecer una normalización de los productos, lo que facilitaría el intercambio y la comercialización entre los distintos países. Entre 2000 y 2001 las normas ISO se aplicaron finalmente al campo sanitario³¹. AENOR es la organización española que se encarga de la traducción y la comercialización de estas normas.

Las normas ISO son genéricas para cualquier tipo de organización. Este hecho supuso una gran dificultad en la interpretación y la adaptación de las normas al ámbito sanitario al comienzo de su uso. Algunas in-

investigaciones³², realizadas sobre la aplicación de las normas ISO en el contexto de la sanidad, muestran que las guías publicadas por los organismos nacionales sobre la aplicación de dichas normas al tratamiento de enfermos agudos tienen unas interpretaciones que, en ocasiones, son ligeramente diferentes y, en algunos casos, radicalmente distintas, lo que cambia completamente el significado de algunos aspectos de la norma. Esta dificultad en la interpretación, junto con el aumento del «papeleo» en la primera fase de su aplicación, han sido las críticas más comunes a este modelo.

Las normas ISO 9000 pueden aplicarse a toda la organización o a un sector específico. Es común su aplicación a departamentos de apoyo (no clínicos), que normalmente son los que más se pueden beneficiar de la estandarización y la reducción de la variabilidad.

La JCAHO fue introducida en España por la Fundación Avedis Donabedian (FAD) entre 1996 y 1997, con el objetivo de implantar un sistema de acreditación muy extendido en los hospitales estadounidenses. La diferencia principal de la JCAHO con las normas ISO o el EFQM es que la JCAHO es un sistema de acreditación específico para organizaciones sanitarias, por lo que las normas se centran en los pacientes y cubren los aspectos principales de la práctica clínica. Los conceptos y el lenguaje empleados por el modelo son más cercanos a los profesionales clínicos, lo que ayuda a motivarlos e incorporarlos en el proceso de mejora de la calidad. La dificultad principal para la implantación de la JCAHO en España puede estribar en la creencia de que estas normas se crearon para el sistema sanitario estadounidense, que es en muchos aspectos diferente del español³³.

La Joint Commission International (JCI) creó su primer programa de acreditación de hospitales³⁴ en 1999, con una flexibilidad que tiene en cuenta las características sociales, políticas, religiosas y económicas en el país en que se aplica. Este nuevo programa acercó la acreditación de la JCAHO al sistema sanitario español. En estos momentos, 25 organizaciones (hospitales, organizaciones sanitarias y centros sociosanitarios) han sido evaluadas por la FAD mediante los estándares de la Joint Commission³⁵, y 5 hospitales han conseguido ya la acreditación³⁶.

Estudio de un caso práctico: la aplicación internacional del Maryland Quality Indicator Project®

En este artículo presentamos un proyecto que empezó hace casi 20 años en Maryland, Estados Unidos. El estudio se centra en el Maryland Quality Indicator Project® (QIP), que es en la actualidad el sistema de medición de rendimiento más utilizado en el mundo. Casi

2.000 participantes de Estados Unidos, Austria, Bélgica, Países Bajos, Alemania, Singapur, Taiwán y Reino Unido se han unido al proyecto aceptando su filosofía y, por tanto, sus métodos de evaluación e interpretación^{37,38}. Una breve descripción del proyecto y su filosofía facilitará la comprensión de algunos conceptos que se han expuesto previamente: indicadores, índices, calidad, rendimiento y responsabilidad.

A veces un título es sólo un título

A mediados de los años ochenta, cuando se estaba diseñando el Maryland QIP, se pensaba que los indicadores podían medir directamente la calidad. El nombre del proyecto, por tanto, incluía la palabra calidad, prometiendo implícitamente a sus usuarios que la calidad puede medirse. Con el tiempo, la experiencia de este proyecto puso de manifiesto que las mediciones, cuando se construyen correctamente y con validez, pueden medir el rendimiento con bastante fiabilidad, pero que la traducción de estas medidas en calidad requiere la aplicación de una estrategia posterior. Esta estrategia se ha basado en la formación de los usuarios, de forma similar al ejemplo del perro de caza y el cazador que presentábamos previamente. El QIP adoptó una doble estrategia para la aplicación óptima de las medidas de rendimiento: producir los mejores perros de caza (indicadores) y entrenar continuamente al cazador (el intérprete de los resultados del indicador)³⁹. Esta doble estrategia no sólo convirtió al QIP en un proyecto de formación en sí mismo, sino que también demostró la dificultad de lograr resultados consistentes si sólo se proporciona a los usuarios los instrumentos de medida o la formación, sin un proyecto para la aplicación práctica de la formación recibida. Teniendo en cuenta estos conceptos, la filosofía del QIP es que los indicadores «deben indicar» en lugar de medir. El QIP, como proyecto, ha mostrado su validez para identificar áreas de dudoso rendimiento y, a la larga, indicar calidad. Consideramos necesario señalar que el título del proyecto, aunque utiliza un término que en sentido estricto puede ser malinterpretado en la actualidad, refleja en realidad las expectativas existentes durante las últimas 2 décadas, mientras que nuestros conocimientos han ido evolucionando en función de las experiencias prácticas.

La interpretación de la calidad es de carácter local

Las medidas y los indicadores utilizados por el QIP no llevan un juicio de valor implícito. Aunque se han ido

incluyendo algunos indicadores estratificados en función de la evidencia (p. ej., anticoagulante, profilaxis tras un infarto de miocardio), el QIP se ha mantenido como un proyecto basado en indicadores y ha mejorado continuamente su habilidad para identificar, interpretar y predecir resultados en función de las características del hospital y los resultados obtenidos previamente. Sin embargo, el punto fuerte del proyecto es el uso de los datos y los resultados obtenidos mediante indicadores como guía para el trabajo en hospitales de forma individual, en el ámbito regional o incluso nacional, traduciendo los cocientes y las tendencias obtenidos en dimensiones de adecuación. Hay abundante bibliografía sobre el proyecto y muchos informes acerca de que las magnitudes obtenidas con un mismo indicador, aparentemente similares, tienen distinto significado según el sistema sanitario y las características culturales de la comunidad⁴⁰⁻⁴³.

Funcionamiento del QIP

Las principales líneas del QIP son: medición, formación, investigación y desarrollo, y responsabilidad. La figura 1 muestra cómo se relacionan entre sí las diferentes líneas y cuáles son los términos clave para cada una de ellas.

Figura 1. Líneas principales del Quality Indicator Project.



Las medidas son principalmente los indicadores que se han ido desarrollando durante los últimos 20 años. Hay casi 50 medidas en total, ya que la mayoría de los indicadores están compuestos por más de una medida. Los conjuntos de indicadores de hospitalización, servicios ambulatorios, tratamiento crónico y psiquiatría son los más utilizados actualmente en todo el mundo. Comenzaron a utilizarse en los años ochenta, y desde entonces se ha ido demostrando su utilidad sanitaria en la bibliografía científica internacional. Los indicadores sobre la seguridad de la atención sanitaria son relativamente nuevos e innovadores, al no haber hasta el momento otro conjunto de indicadores sobre esta cuestión (incluida la seguridad del uso del medicamento) que midan sistemáticamente la práctica clínica, evalúen su nivel de seguridad y difundan modelos de mejores prácticas entre los participantes. El estudio piloto de los indicadores de seguridad de la asistencia sanitaria ya ha sido realizado en Singapur y Australia.

Las unidades de análisis también varían con medidas de este tipo. La unidad de análisis principal para los grupos de medidas más tradicionales (hospitalización, servicios ambulatorios, tratamiento crónico y psiquiatría) es la institución. Los datos de pacientes no se analizan de forma individual, sino agregados para crear una estadística global del servicio, departamento u hospital. Al no utilizar datos individuales de pacientes como una unidad de análisis, se incrementa la confianza de los participantes en la confidencialidad de la historia clínica y la información médica. Sin embargo, en el caso de los indicadores de seguridad de la atención sanitaria, la unidad de análisis es más variable. Además del departamento, el servicio y el proceso, hay ocasiones en las que es necesario recoger información individual del paciente, aunque no sea necesario introducirla en la base de datos del proyecto. La información individual es útil porque es necesario identificar ciertos errores en el hospital para realizar una revisión interna, y si estos errores se producen con poca frecuencia, será difícil agregar la información individual y convertirla en un cociente. En cualquier caso, independientemente de la unidad de análisis utilizada, los informes proporcionados por el proyecto son siempre anónimos, y el análisis de información individual de los pacientes no está entre los objetivos del QIP.

La línea de formación del proyecto es uno de los factores distintivos respecto a otras iniciativas. Al ser una colaboración de entidades geográficamente dispersas, conlleva un componente de formación que es útil tanto para los participantes como para el propio personal del proyecto. La formación está dirigida a los distintos grupos profesionales de las instituciones sanitarias: administración, gestión, profesionales sanitarios, profesionales dedicados a la mejora de la calidad, y expertos en gestión de historias clínicas y gestión de información. La formación del QIP es muy dinámica y está

en constante desarrollo mediante los programas de formación continuada, formación de formadores y seminarios de formación periódicos, impartidos internacionalmente.

La línea de investigación y desarrollo del QIP también lo distinguen de otras iniciativas similares, ya que el QIP continúa siendo un proyecto de investigación que recibe un *feedback* constante de los participantes, útil para realizar las modificaciones necesarias, e incrementa así constantemente su valor. La línea de investigación y desarrollo del QIP también proporciona la oportunidad de experimentar con indicadores a partir de la participación voluntaria de instituciones de todo el mundo. A través de esta colaboración voluntaria, se ha desarrollado un nuevo grupo de indicadores innovadores de I+D, cuya validez ha sido estudiada durante 3 años.

La línea de responsabilidad del QIP está dirigida a demostrar a los diferentes colectivos la utilidad que el proyecto tiene para ellos. Hay diferentes grupos de interés ante los cuales el QIP debe demostrar su responsabilidad, donde el grupo principal es la propia organización. No obstante, el creciente interés de varios sistemas sanitarios por conocer el impacto que la mejora del rendimiento supone en el acceso a los servicios, el coste de la provisión de servicios y la seguridad del medio en que son proporcionados está sumando nuevos colectivos al público inicial. Por ejemplo, los ministerios de sanidad de Singapur y Austria han colaborado con las instituciones participantes, facilitando el proceso y respetando los objetivos del proyecto de no divulgar públicamente la información que genera. Por tanto, la responsabilidad puede alcanzarse si todos los usuarios están conformes con los beneficios y las limitaciones que un proyecto basado en indicadores puede tener de cara a establecer un mejor sistema sanitario.

¿Por qué el QIP alcanza tal aceptación mundial?

El interés global de los participantes en el proyecto es la búsqueda de sistemas de medición del rendimiento que proporcionen una guía válida y fiable con un nivel de esfuerzo razonable. Independientemente del sistema sanitario o el continente al que pertenezca una organización, el requisito de mejorar el rendimiento y demostrar que dicha mejora se ha realizado se está convirtiendo en una necesidad en todo el planeta. Los argumentos para adoptar o adaptar un proyecto o una metodología específica han ido siempre unidos a una definición de calidad supuestamente diferente. Mientras que la mayoría de los profesionales sanitarios cree que la medición puede ser genérica, suele discrepar en lo que dicha medición significa en términos de rendimiento

bueno, malo o indeseable. Quizá el factor decisivo de estas preguntas sea las expectativas, en lugar de la definición de calidad. Si se calibran y se normalizan las expectativas, se descubre que la definición de calidad es mucho más uniforme de lo que parecía en un principio. De hecho, en el trabajo de Donabedian encontramos las dimensiones operativas de calidad: adecuada, responsable, adaptable, aceptable y económicamente rentable. En nuestra opinión, un conjunto similar de dimensiones puede aplicarse a las medidas de rendimiento, que deberían ser: válidas, orientadas a resultados, responsables, pertinentes y relativas. De esta forma, los indicadores deben formar parte de un sistema de medición con objetivos bien definidos (válido), que puede cuantificar rendimiento o resultados (orientado a resultados), promueve la responsabilidad con respecto a la actuación (responsable), se mantiene actualizado con los avances científicos del sector (pertinente), y tiene en cuenta las características del contexto y el ambiente en que se aplicará (relativo). Éste es el atractivo del QIP, y el motivo de su éxito en Estados Unidos, Europa y Asia ha sido introducir una potente metodología para la medida de la actuación, teniendo en cuenta las costumbres y las expectativas locales a la hora de transformar los resultados obtenidos mediante indicadores en las 5 dimensiones de calidad.

Ejemplos recientes de aplicación internacional del QIP

En un artículo reciente⁴⁰ se demuestran los perfiles de uso del QIP y su aplicación en Estados Unidos, Europa y Asia. Este artículo proporciona una información única, ya que ningún otro proyecto tiene la envergadura ni la trayectoria histórica del QIP. En la tabla 1 se muestra cómo los indicadores seleccionados para ser utilizados en hospitales de Asia, Europa y Estados Unidos presentan similitudes significativas en cuanto a magnitud y distribución. Los datos de la tabla 1 sugieren que, a pesar de la variabilidad en la práctica asistencial entre los distintos sistemas sanitarios, la frecuencia real con la que se proporciona un cierto tipo de servicio o la magnitud de las prácticas evaluadas mediante indicadores no difiere significativamente entre los distintos sistemas sanitarios. La recomendación del artículo es continuar observando y comparando dichos perfiles a largo plazo y establecer, quizá por vez primera, una línea de base real para los cocientes de frecuencia, prevalencia e incidencia reflejados en los indicadores del QIP. En estos momentos se está realizando un segundo estudio con el objetivo de actualizar los datos estadísticos y determinar si son comparables durante varios años.

Tabla 1. Tasas medias ponderadas para Estados Unidos, Asia oriental y Europa

Descripción de la medida	Estados Unidos			Europa			Asia oriental		
	1999	2001	Variación (%)	1999	2001	Variación (%)	1999	2001	Variación (%)
Mortalidad intrahospitalaria total	2,74	2,57	-6,20	2,31	1,91	-17,32	1,65	1,78	7,88
Mortalidad perioperatoria estratificada por clasificación ASA	0,34	0,34	0,00	0,18	0,14	-22,22	0,11	0,17	54,55
Cesáreas primarias	15,47	16,95	9,57	13,19	19,37	46,85	21,61	22,95	6,20
Cesáreas repetidas	70,41	78,66	11,72	60,31	73,07	21,16	75,07	92,01	22,57
Frecuencia total de cesáreas	22,14	24,67	11,43	20,24	24,35	20,31	31,63	31,9	0,85
Parto vaginal tras cesáreas	29,33	21,33	-27,28	26,38	26,95	2,16	7,12	8,18	14,89
Reingresos no programados hasta 31 días tras el alta por el mismo proceso patológico u otro relacionado	4,39	4,37	-0,46	2,00	1,52	-24,00	4,69	4,39	-6,40
Retornos no programados a la unidad de cuidados intensivos	3,08	3,24	5,19	5,85	4,08	-30,26	6,22	4,77	-23,31
Reintervenciones quirúrgicas no programadas	1,41	1,37	-2,84	0,75	1,16	54,67	0,4	0,65	62,50
Reingresos no programados dentro de las 72 h tras el alta	1,74	1,9	9,20	1,85	5,47	195,68	2,11	2,18	3,32

Reproducido con autorización de Kazandjian et al⁴⁰.

Comentarios finales

El análisis sobre la medida sistemática del rendimiento a través de indicadores está en consonancia con las iniciativas que se están llevando a cabo en la sanidad española. Mientras que la organización y la responsabilidad de la provisión de servicios sanitarios sigue filosofías nacionales y regionales, los objetivos de ambos modelos son el mismo: la comprensión de la actuación, la responsabilidad de cara a los profesionales y los receptores de los servicios sanitarios, y la mejora de la eficiencia en la producción de servicios más seguros. El interés en el tema de la medición de resultados y el nivel de las iniciativas llevadas a cabo para ello en España son elevados, como cabe esperar en estadios iniciales de desarrollo de modelos nacionales, a la vez que parece haber una gran variabilidad entre los distintos proyectos, las comunidades autónomas y, en algunas ocasiones, las filosofías de base. El modelo EFQM ha suscitado mucho interés y atención en el sistema sanitario de nuestro país, igual que en muchos otros países europeos. La adaptación de sistemas de evaluación para la acreditación requiere un grupo de participantes con mucha formación, que no sólo estén interesados en tipos de responsabilidad y mejora de calidad, sino que también sean conscientes del esfuerzo necesario para construir el sistema, evaluarlo y comprobar su validez. Podemos considerar a España un modelo significativo entre muchos países europeos en su deseo de seguir varias iniciativas simultáneas para conseguir las mismas metas de calidad y responsabilidad. Siguiendo varias líneas de actuación (medir el rendi-

miento mediante indicadores, crear sistemas de acreditación, potenciar la responsabilidad con la sociedad mediante rankings de hospitales, evaluar la satisfacción y las expectativas de los usuarios de los servicios y considerar la posibilidad de realizar técnicas de *benchmarking* internacional), los responsables de la toma de decisiones necesitan sopesar el coste de los programas teniendo en cuenta los resultados esperados. En este artículo se propone que la evaluación del rendimiento mediante indicadores puede servir de base (como en el caso del Maryland Quality Indicator Project[®]) para la discusión sobre los modelos de actuación aceptables, y a partir de ahí identificar objetivos y modelos de comparación. También se propone que el alto nivel de información del sistema sanitario español actual constituye un escenario en el que la introducción de nuevas ideas probablemente encontraría una buena aceptación y una correcta aplicación. Por tanto, corresponde a los responsables de plantear nuevas ideas y recomendaciones al sistema sanitario español evaluar cuidadosamente el impacto que éstas pueden tener en el sistema actual, y a partir de ahí formar progresivamente a los usuarios sobre los retos, los beneficios y los modos razonables de interpretar los resultados obtenidos. De cara al futuro, los autores de este artículo consideran necesaria, aunque no suficiente, la aplicación de sistemas de medida basados en la evidencia (como indicadores bien definidos) para la medida continua y sistemática de los modelos de actuación sanitaria, de forma individual, organizacional, regional e incluso nacional. Asimismo, se considera que la responsabilidad sólo podrá lograrse con la aplicación de dichas medidas, de forma rigurosa y sostenible.

Bibliografía

1. O'Leary D. JCAH plans new series of quality indicators based on outcome, clinical standards. *Rev Fed Am Health Syst* 1986;19:26-7.
2. O'Leary DS. The JCAH quality initiative: what can hospitals and physicians expect? Interview by Emily Friedman. *Physician Exec* 1987;13:2-6.
3. Collopy BT. Developing clinical indicators: the ACHS Care Evaluation Program. *Aust Clin Rev* 1990;10:83-5.
4. Thomson RG, McElroy H, Kazandjian VA. Maryland Hospital Quality Indicator Project in the United Kingdom: an approach for promoting continuous quality improvement. *Qual Health Care* Mar 1997;6:49-55.
5. Matthes N, Wiest A. Qualität im Krankenhaus-Überlegungen zu Begriffen und Methoden der Leistungsbewertung. En: Arnold M, Klauber J, Schellschmidt H, editors. *Krankenhaus-Report 2002*. Stuttgart/New York: Schattauer; 2003. p. 161-74.
6. Klazinga N. Concerted action programme on quality assurance in hospitals 1990-1993 (COMAC/HSR/QA). Global results of the evaluation. *Int J Qual Health Care* 1994;6:219-30.
7. Casparie AF, Klazinga NS, Van Everdingen JJ, Touw PP. Health-care providers resolve clinical controversies: the Dutch consensus approach. *Aust Clin Rev* 1987;7:43-7.
8. Mainz J, Bartels P, Rhode P. Accountability and Measurement in the Danish Healthcare System. En: Kazandjian VA, editor. *Accountability through measurement: a global healthcare imperative*. Milwaukee: ASQ Quality Press; 2002. p. 204-16.
9. Kazandjian VA. *Accountability through measurement: a global healthcare imperative*. Milwaukee: ASQ Quality Press; 2002.
10. WHO Regional Office for Europe. Measuring hospital performance to improve the quality of care in Europe: a need for clarifying the concepts and defining the main dimensions. Report on a WHO workshop. Barcelona, 10-11 de enero de 2003.
11. Wennberg JE. Dealing with medical practice variations: a proposal for action. *Health Aff (Millwood)* 1984;3:6-32.
12. Kazandjian VA. Performance indicators: pointer dogs in disguise –a commentary. *J Am Med Rec Assoc* 1991;62:34-6.
13. Murray CJ. *Development data constraints and the Human Development Index*. Geneva: UNRISD; 1991.
14. Krottscheck C, Narodslawsky M. The Sustainable Process Index: a new dimension in ecological evaluation. *Ecological Engineering* 1994;241-58.
15. Desai M. Income and alternative measures of well-being. En: Westendorff DG, Ghai D, editors. *Monitoring social progress in the 1990s*. Brookfield: Averbury, published for UNRISD; 1993. p. 23-40.
16. Daly HE, Cobb JB, Cobb CW. *For the common good: redirecting the economy toward community, the environment, and a sustainable future*. 2nd ed. Boston: Beacon Press; 1994.
17. Repullo R. Los retos para promover la calidad en el nuevo escenario descentralizado del Sistema Nacional de Salud. *Rev Calidad Asistencial* 2002;17:199-200.
18. Programa de Acreditación de Centros. *Manual de Estándares (ME 2 1_01)*. Fundación Agencia de Calidad Sanitaria de Andalucía. Sevilla: Consejería de Salud de la Junta de Andalucía, 2002.
19. Arcelay A. Adaptación de un modelo de gestión de calidad total al sector sanitario. *Rev Calidad Asistencial* 2000;115:184-91.
20. Guix J, Simón R, Nualart L, Surroca RM, Carbonell JM. Utilización de modelos como herramienta de diagnóstico y mejora de la calidad: EFQM y Joint Commission. *Rev Calidad Asistencial* 2001;16:308-12.
21. Guía para autoevaluación de Centros Sanitarios utilizando como referencia el modelo de autoevaluación de la European Foundation for Quality Management (EFQM). MSD; 2001.
22. *Manual de evaluación y mejora de la satisfacción de las personas en las organizaciones de servicios*. Vitoria: Osakidetza-Servicio Vasco de Salud; 2001.
23. Identificación de un conjunto de indicadores claves de calidad en el ámbito hospitalario según el Modelo Europeo de Excelencia (Modelo EFQM de Excelencia). Club Gestión de Calidad; 2001.
24. Ochoa AO, Torrijos GP, Sánchez DS. Cuadro de mandos integral a nivel hospitalario basado en indicadores del modelo EFQM de excelencia. *Gest Hosp* 2002;13(1):9-25.
25. Cortes JA, et al. Análisis del liderazgo en atención primaria mediante el modelo EFQM, con motivo de la realización de un plan estratégico. *Cuadernos de Gestión* 2001;7:119-32.
26. Ramón C, Ricci C. Proceso de autoevaluación de la Fundación Hospital Manacor mediante el Modelo Europeo de Gestión de Calidad Total. *Rev Calidad Asistencial* 2000;15:151-4.
27. Fernández-León A, Berros MA, García M. La experiencia del Hospital Monte Naranco en la Aplicación del Modelo de la EFQM. *Rev Calidad Asistencial* 2000;15:155-60.
28. Lloria P, Zarzo A, Ramos J. Introducción a la implantación de la EFQM en el Hospital Universitario La Fe. *Rev Calidad Asistencial* 2000;15:135-41.
29. Moracho O, Colina A, Amondarain MA, Aguirre L, Ruiz-Álvarez E, Salgado MV. Experiencia práctica del proceso de evaluación externa con el Modelo de Excelencia de la EFQM en el Hospital de Zumárraga. *Rev Calidad Asistencial* 2001;16:322-9.
30. Arcelay A, Sánchez E, Hernández L, Inclán G, Bacigalupe M, Letona J, et al. Self Assessment of all the health centres of a public health service through the European Model of Total Quality Management. *Int J Health Care* 1999;12:54-8.
31. Orbea T. Aportación de las ISO 9000 a los servicios sanitarios. *Rev Calidad Asistencial* 2001;16:484-5.
32. Sweeney J, Heaton C. Interpretations and variations of ISO 9000 in acute health care. *Int J Qual Health Care* 2000;12:203-9.
33. Bohigas L. Accreditation across borders: The introduction of Joint Commission Accreditation in Spain. *Jt Comm J Qual Improv* 1998;24:226-31.
34. Joint Commission International Accreditation. *Estándares internacionales de acreditación de hospitales*. Joint Commission on Accreditation of Healthcare Organizations; 2000.
35. Bañeres J. *Sistemas Internacionales de evaluación externa*. *Rev Calid Asist* 2003;18(2):69-71.
36. Joint Commission on Accreditation of Healthcare Organizations. Disponible en: <http://www.jcaho.org/>
37. Center for Performance Sciences. Disponible en: <http://www.cpsciences.com/>
38. International Quality Indicator Project. Disponible en: <http://www.internationalqip.com/>
39. Kazandjian V. Public disclosure: A hospital quality improvement project's perspective. En: Goldfield N, editor. *Physician profiling and risk adjustment*. 2nd ed. Gaithersburg: Aspen Publishers, Inc.; 1999. p. 95-106.

40. Kazandjian VA, Matthes N, Wicker K. Are performance indicators generic? The international experience of the Quality Indicator Project® [en prensa]. *J Eval Clin Pract* 2003;9:265-76.
 41. Kazandjian VA, Lawthers J, Cernak CM, Pipes FC. Relating outcomes to processes of care: the Maryland Hospital Association's Quality Indicator Project (QI Project). *Jt Comm J Qual Improv* 1993;19:530-8.
 42. Kazandjian VA, Summer SJ, Kazandjian JA. Evaluating the appropriateness of care: a study of cesarean section rates. *QRB Qual Rev Bull* 1989;15:206-14.
 43. Kazandjian VA, Thomson RG, Law WR, Waldron K. Do performance indicators make a difference? *Jt Comm J Qual Improv* 1996;22:482-91.
-