

La invasión de los indicadores compuestos. Riesgos y beneficios para la gestión de la calidad

Pedro J. Saturno

Unidad de Medicina Preventiva y Salud Pública. Universidad de Murcia. Murcia. España.

Correspondencia: Dr. Pedro J. Saturno.
Facultad de Medicina. Universidad de Murcia.
Campus Universitario de Espinardo.
30100 Murcia. España.
Correo electrónico: psaturno@um.es

Resumen

Los indicadores compuestos son un tipo particular de indicadores que resumen en un solo número la medición de varios subindicadores. Su gran atractivo y potencial idoneidad para la comparación de centros, sistemas y países están haciendo que se les preste cada vez más atención. Sin embargo, su utilización no está exenta de riesgos metodológicos derivados de las características de los distintos subtipos de indicadores compuestos existentes, cada uno con sus ventajas e inconvenientes. Este artículo revisa y clasifica los diversos tipos de indicadores compuestos de que podemos disponer en la actualidad y su aplicación a la gestión de la calidad en servicios de salud.

Palabras clave: Indicadores compuestos. Control de calidad. Servicios de salud.

Introducción

En la gestión de la calidad se entiende por monitorización la medición sistemática y planificada de indicadores de calidad¹, una actividad conducente a controlar que se está en niveles preestablecidos y que tiene como objetivo identificar la existencia o no de situaciones problemáticas que hay que evaluar o en las que hay que intervenir. Los indicadores a monitorizar pueden derivarse de ciclos de mejora, de actividades de diseño de los servicios, o ser fruto de una selección de aspectos o servicios relevantes de un determinado centro cuya calidad interesa controlar. Sea cual sea el origen de la monitorización, hay que utilizar buenos indicadores, es decir, indicadores válidos, fiables y apropiados, características cuyo significado práctico ha sido descrito en otro lugar¹. Sin embargo, además de estos requisitos comunes que deben tener todos los indicadores, existen otras peculiaridades que distinguen tipos diferentes de indicadores y tienen gran trascendencia metodológica para la práctica de la monitorización.

Los indicadores compuestos (*composite indicators*) son uno de los tipos de indicadores que pueden identificarse en función de estas peculiaridades con trascendencia práctica. Su característica principal es que resumen o agregan varios

Abstract

Composite indicators are a particular type of indicators that summarize in a single number the measurements of several sub-indicators. They are particularly attractive and potentially adequate for comparing centres, systems or countries; therefore, composite indicators are increasingly considered and used. However, the construction and utilization of the different types of composite indicators risk some methodological pitfalls, along with specific advantages and disadvantages. This article is a review of the characteristics of the different types of composite indicators, and their use for quality management in health services.

Key words: Composite indicators. Quality of care monitoring. Health services.

indicadores específicos en un solo número. Que se los conciba como número-resumen los hace especialmente útiles como indicadores para comparar centros, áreas o sistemas, aunque también pueden utilizarse para unidades organizativas más simples. La atención notablemente creciente que se les viene prestando en todos los ámbitos, sobre todo internacionales, sanitarios y no sanitarios²⁻⁴, junto con la existencia de varios y bien definidos subtipos, con sus ventajas e inconvenientes, beneficios y riesgos, hace que merezca la pena que se les preste una atención especial.

Este artículo revisa las características de los diversos tipos de indicadores compuestos, con una consideración específica a los ejemplos de su aplicación en los servicios de salud.

Indicadores compuestos: concepto y relevancia

Los indicadores compuestos resumen, como se ha dicho, en un solo número las mediciones de 2 o más indicadores, que en este artículo se denominarán subindicadores. Para algunos autores sería deseable que los subindicadores estuviesen relacionados entre sí, incluso estadísticamente⁵; sin embargo, el principal cuerpo teórico en este campo considera que los subindicadores no tienen que estar necesaria-

mente relacionados, ni tener una unidad de medida común, ni tampoco tener cada uno de ellos una importancia relativa obvia que pueda servir de guía en una eventual ponderación de su agrupación en un solo indicador compuesto.

La atención que se les está prestando a los indicadores compuestos es relativamente reciente, pero en poco tiempo se han situado en el centro del debate teórico-práctico, tanto en el campo social y económico como en el sanitario. En el campo social y económico, las razones para este interés, según se resumen en el documento base de una reunión consultiva de la Unión Europea sobre este tema, celebrada en Bruselas en marzo de 2002, son⁴:

- Pueden ser utilizados para resumir aspectos o problemas complejos y multidimensionales, y ayudar en la toma de decisiones.

- Proveen una visión global, y pueden ser interpretados más fácilmente que el escrutinio de muchos indicadores por separado, facilitando los *rankings* de países y regiones en cuestiones complejas.

- Son generalmente mejor entendidos y atractivos para el gran público, al proporcionar una cifra resumen para comparar sistemas y países.

- Pueden ayudar a reducir el tamaño de los listados de indicadores, o a la inclusión de más información dentro de unos límites de espacio establecidos.

También se advierte de posibles factores en contra, como que resulten demasiado simples, no bien construidos y mal interpretados, así como la falta de transparencia en las valoraciones y decisiones adoptadas en cuanto a la selección de los subindicadores, la metodología de agrupación, y que, a

pesar de todo, pueda necesitarse mayor cantidad de datos para su construcción, puesto que hay que medir todos los subindicadores que se incluyan.

Sin embargo, todas estas circunstancias potencialmente negativas se consideran retos metodológicos que hay que solucionar, antes que inconvenientes reales que frenen una tendencia y un interés en este tipo de indicadores, que va muy visiblemente en aumento. Sirva como ejemplo la iniciativa del Joint Research Centre (JRC) de la Comisión Europea y la Organización para la Cooperación y el Desarrollo (OCDE), cuyo primer taller sobre indicadores compuestos en el desempeño económico y social de los países, celebrado en mayo de 2003, se presenta como un verdadero hito en el que se revisó lo realizado hasta ese momento y se propusieron actividades de continuidad sobre el mismo tema. Tiene especial interés el documento elaborado por el Grupo de Estadística Aplicada del JRC sobre la situación actual de la metodología y práctica para el desarrollo de indicadores compuestos⁴.

Uno de los indicadores compuestos más referido en el campo económico y social es el Human Development Index (Índice de Desarrollo Humano), un indicador utilizado por Naciones Unidas que agrega la esperanza de vida, las tasas de escolaridad y analfabetismo y el producto interior bruto (PIB) per cápita. En el campo sanitario es también un organismo de Naciones Unidas, la Organización Mundial de la Salud (OMS), el que construyó y utilizó el Informe de Salud Mundial del año 2000⁶, un indicador compuesto que agrupó 5 componentes para clasificar los sistemas de salud de 191 países y que ha sido discutido desde muchos puntos de vista, pero que también ha contribuido a poner el foco en este tipo de indicadores para seguir perfeccionándolos. La tabla 1 contiene otros ejemplos destacados de utilización de indicadores compuestos en sistemas de salud.

Tabla 1. Algunos ejemplos de indicadores compuestos en sistemas de salud

Canadá: combinación de 15 indicadores (relativos a atención prenatal, salud comunitaria, servicios para los ancianos, eficiencias, recursos) agrupados y ponderados en función de 6 categorías, para cada una de las 54 regiones sanitarias³.

Estados Unidos: *ranking* de estados sobre la calidad de atención a pacientes de Medicare³: incluye 22 indicadores (en atención a infarto, accidente cerebrovascular, insuficiencia cardíaca, neumonía, cáncer de mama, diabetes). Cada estado tiene una posición en la lista de cada indicador. El indicador resumen para cada estado es la suma, o el promedio, de las posiciones en los 22 listados.

Estados Unidos: construcción de 4 indicadores compuestos (atención en la consulta, calidad clínica, accesibilidad y servicio al cliente, y vacunaciones) combinando 23 subindicadores (9 del Consumer Assessment of Health Plan Study y 14 del Health Plan Employer Data Information System)⁵.

España: Top 20. Clasificación de los hospitales sumando su posición en el *ranking* de 6 indicadores (mortalidad, complicaciones, estancia media, índice de ocupación, coeficiente de ambulatorización, coste por unidad de producción). Un acercamiento parecido al *ranking* 100 Top Hospitals de Estados Unidos.

Reino Unido: indicadores compuestos de alto nivel del National Health Service. Tasas medias ponderadas de varios indicadores relacionados (tabla 5)¹³.

Reino Unido: indicador del King's Fund para comparar las 120 Health Authorities de Inglaterra, Escocia y País de Gales³, compuesto de 6 subindicadores ponderados (muertes por cáncer, muertes por enfermedad cardíaca, número de personas en lista de espera, número de personas en lista de espera más de 12 meses, número de operaciones de cadera, muertes prematuras y evitables).

Reino Unido: indicadores de calidad para los servicios de atención primaria, en el nuevo contrato con los médicos generales del National Health Service¹⁵; 146 indicadores que se resumen con un sistema de puntuación (un total máximo de 1.050 puntos, distribuidos en 7 grupos de indicadores, cada uno a su vez resumido en la puntuación que le corresponde)

Organización Mundial de la Salud: indicador compuesto (que reúne resultados en salud, desigualdades en salud, respuesta del sistema de salud, desigualdades en la respuesta del sistema, financiación equilibrada) para clasificar los sistemas de 191 países (World Health Report 2000)⁶

Diversos tipos de indicadores compuestos según su metodología de construcción

Como consecuencia del número y trascendencia de las iniciativas internacionales en relación con los indicadores compuestos, existe en la actualidad un cierto cuerpo de doctrina en la metodología para su construcción, y se han identificado aquellos pasos o decisiones que pueden resultar especialmente problemáticos. La tabla 2 resume el esquema metodológico general y las principales consideraciones para cada paso.

De todas las decisiones metodológicas que se mencionan en la tabla 2, la mayoría de las discusiones se han venido centrando en las correspondientes al paso 2 (selección de subindicadores), al 4 (valoración de la relación existente entre los subindicadores) y, sobre todo, al 5 (normalización, ponderación y agregación de los subindicadores). Sobre la valoración de la relación entre los subindicadores ya hemos mencionado que existen 2 posiciones contrapuestas: considerar que es necesario que estén relacionados para que el indicador compuesto sea consistente y, al contrario, considerar que un indicador compuesto puede resumir varios y diferentes aspectos de un fenómeno no necesariamente relacionados entre sí, aunque todos tengan el nexo del fenómeno general que contribuyen a medir. La mayoría de los indicadores compuestos utilizados parece que asumen esta segunda aproximación, aunque también se utilizan con cierta frecuencia técnicas como el análisis factorial o el análisis de componentes principales para agrupar estadísticamente los subindicadores en grupos, factores o conglomerados que funcionan entonces como variables para la agregación en un solo indicador compuesto, incluso con la ponderación que pueda resultar de este análisis estadístico. Todo ello no obvia, sino que complementa, la importantísima consideración inicial sobre la adecuación y representatividad de los subindicadores seleccionados con relación al fenómeno a medir por el indicador compuesto.

En función de todas estas características, pueden identificarse diferentes tipos de indicadores compuestos, con métodos de elaboración más o menos complejos, según describe de forma resumida la tabla 3. En los apartados siguientes se resume lo más destacado de cada uno de los 3 grandes grupos (sin ponderación y con ponderación, y dentro de los primeros sin reconversión y con reconversión de las escalas de medida de los subindicadores), subrayando los que son potencialmente más fáciles de aplicar y entender, y de más utilidad para la mejora continua, así como los ejemplos existentes en el ámbito de los servicios de salud.

Indicadores compuestos sin reconversión de escala ni ponderación de los subindicadores

Son los más simples de calcular, y por tanto posiblemente también de entender, consisten esencialmente en la medición de los subindicadores y el establecimiento de una forma concreta de agregar los resultados en un solo número, que es el indicador compuesto. El método de agregación es lo que les hace diferentes entre sí, y es también el origen de las ventajas e inconvenientes de cada uno.

La suma o el promedio de los *rankings* obtenidos en cada subindicador para comparar un número o “*ranking-resumen*” es el método que se utilizó en Estados Unidos para comparar los diferentes estados con relación a la calidad de la atención a pacientes de Medicare³ (tabla 1) y que utiliza la clasificación Top 20 de los hospitales en España⁷. Tiene como problema que puede producir valores medios que enmascaran situaciones de subindicadores con niveles muy bajos compensados por otros con niveles altos. Problema semejante al que tiene la media de cumplimientos de los subindicadores como indicador compuesto, método que utilizaba el extinto Instituto Nacional de la Salud (INSALUD) para resumir la valoración de las llamadas Normas Técnicas Mínimas (en realidad una serie de criterios o subindicadores) en atención primaria^{8,9}.

Tabla 2. Esquema general para la construcción de indicadores compuestos. Pasos y principales consideraciones metodológicas

| Paso | Considerar |
|--|---|
| 1. Decidir sobre el fenómeno, aspecto o problema a medir | Hasta qué punto se beneficiaría de un indicador compuesto |
| 2. Selección de subindicadores | Justificación de la selección: relevancia y validez de cada indicador a incluir. No hay ninguna fórmula inequívoca para seleccionar qué y cuántos subindicadores |
| 3. Valorar la existencia y calidad de los datos para la medición de los subindicadores seleccionados | Métodos para tratar la ausencia parcial de datos (<i>missing</i>) para algún subindicador |
| 4. Valorar la relación existente entre los subindicadores | Pueden ser de utilidad métodos como el análisis de componentes principales, factorial y de correlación en general |
| 5. Normalización, ponderación y agregación de los subindicadores | Justificar la decisión sobre cuál de los múltiples métodos posibles sería el más adecuado |
| 6. Comprobar la solidez y estabilidad de los resultados | Realizar análisis de sensibilidad a los cambios en la composición del indicador (p. ej., subindicadores que se incluyen, ausencia parcial de datos, método de ponderación) y expresar resultados con límites de confianza |

Fuente: Elaboración propia a partir del informe del Applied Statistics Group⁴.

Tabla 3. Indicadores compuestos: tipos y métodos de agregación

| Tipo | Método |
|--|---|
| 1. Sin ponderación de los subindicadores ^a | |
| 1.1. Sin reconversión de escalas de medida de los subindicadores | |
| 1.1.1. Suma de <i>rankings</i> | $\sum_{i=1}^N \text{Rank}_i$ |
| 1.1.2. <i>Ranking</i> medio | $\frac{\sum_{i=1}^N \text{Rank}_i / N}{\Sigma \text{cumplimientos de los subindicadores}} \times 100$ |
| 1.1.3. Porcentaje simple | $\frac{\Sigma \text{cumplimientos de los subindicadores}}{\Sigma \text{de ocasiones en las que se evalúa cada subindicador}} \times 100$ |
| 1.1.4. Media de cumplimientos | $\frac{\Sigma \text{cumplimientos de los subindicadores}}{\Sigma \text{de ocasiones en las que se evalúa cada subindicador}} \times 100$ |
| 1.1.5. Porcentaje de 100% de cumplimientos | $\frac{\text{casos en los que se cumplen todos los subindicadores}}{N.º \text{ de subindicadores evaluados}} \times 100$ |
| 1.1.6. Suma de número de subindicadores por encima de la media menos número de subindicadores por debajo de la media | $\sum_{i=1}^N \text{sgn}[x_i - \bar{x}_i]$ |
| 1.2. Con reconversión de escalas de medida de los subindicadores | |
| 1.2.1. Suma de razones o de proporciones de las diferencias con la media | $\sum_{i=1}^N \frac{x_i - \bar{x}_i}{\bar{x}_i} \text{ o } \sum_{i=1}^N \frac{x_i - \bar{x}_i}{\bar{x}_i}$ |
| 1.2.2. Suma o media de los valores (absolutos o diferencias con medias) estandarizados | $\sum_{i=1}^N \frac{ x_i - \bar{x}_i }{\sigma_{x_i}} \text{ o } \sum_{i=1}^N \frac{x_i}{\sigma_{x_i}}$ solos o promediado a Σ en numerador y N en denominador |
| 1.2.3. Suma de valores reconvertidos en función de rango | $\sum_{i=1}^N \frac{x_i - \min(x_i)}{\text{rango}(x_i)}$ |
| 2. Con ponderación de los subindicadores | |
| 2.1. Análisis de componentes principales y análisis factorial | Combinaciones estadísticas de grupos de subindicadores |
| 2.2. Regresión múltiple | Modelo de regresión múltiple con los subindicadores como variables independientes y un indicador de referencia como dependiente |
| 2.3. Neutralización del efecto de las correlaciones entre subindicadores | $\Sigma [2 \cdot (1 + r_{x_i y_i})] - 1/2 (x_i + y_i)$ |
| 2.4. Distancia a frontera de eficiencia | Ponderación en función de los valores <i>benchmark</i> de la combinación de subindicadores |
| 2.5. Distancia a valores estándar | Ponderación en función de los valores-objetivo o estándar de cada subindicador |
| 2.6. Opinión de expertos (distribución de presupuesto imaginario) | Ponderación según cantidad otorgada a cada subindicador |
| 2.7. Opinión de la población | Ponderación según valor otorgado por consultas estructuradas a la población o sus representantes |
| 2.8. Proceso Analítico Jerárquico (AHP) | Utilización de métodos cuantitativos y cualitativos en comparaciones sucesivas por pares de subindicadores |

N: número total de subindicadores; i: serie de subindicadores; x: resultado de la medición de un subindicador; sgn: signos; $r_{x_i y_i}$: coeficiente correlación entre 2 subindicadores. ^aEste grupo de métodos no implica necesariamente ponderación, pero también pueden ser utilizados con ponderación previa de los subindicadores, en combinación con alguno de los métodos listados en el apartado 2 de esta tabla.

El indicador compuesto que indica el porcentaje de casos que cumplen todos los subindicadores (porcentaje de 100% de cumplimientos) tiene el inconveniente de ser tremendamente restrictivo. Un solo subindicador con niveles de cumplimiento bajos condiciona una impresión general de mala calidad, con independencia de los niveles que puedan tener los otros subindicadores.

Mención aparte dentro de este grupo merecen los indicadores compuestos tipo porcentaje simple, que son probablemente los de más atractivo y utilidad potencial, además de ser, efectivamente, los más utilizados en la práctica, al menos en el ámbito de las encuestas de calidad percibida. Los indicadores compuestos tipo porcentaje simple están construidos para responder a la pregunta: ¿con qué frecuencia han recibido los pacientes atención de calidad?; pregunta que se puede referir a diferentes ámbitos dentro del sistema (servicios, centros, áreas, regiones, sistemas de salud) y unidades de estudio (global, determinados procesos, pacientes, aspectos, etc.).

Su principal peculiaridad es que cada caso evaluado proporciona múltiples "subcasos", tantos como subindicadores sea pertinente evaluar en cada caso concreto, con lo cual se necesitan muestras más pequeñas para contestar con precisión a esa pregunta general ("atención de calidad" o "sin problemas") que las que se necesitarían para evaluar correctamente cada uno de los subindicadores. La tabla 4 contiene un ejemplo clínico en este sentido. Como puede verse, aunque son sólo 15 los casos evaluados, para el indicador compuesto es como si fueran 40, que son las veces en que ha sido pertinente evaluar los subindicadores, considerando cada una de ellas como una "oportunidad" diferenciada para actuar con calidad o no. El indicador compuesto de este ejemplo nos dice que se ha dado atención de calidad en el 70% de las ocasiones.

De forma semejante, los resultados de las encuestas de calidad percibida pueden resumirse con las llamadas "tasas de problemas", que en realidad son indicadores compuestos tipo porcentaje simple. Las "tasas de problemas" resumen las respuestas de varias preguntas sobre diversos aspectos de una variable compleja o dimensión indagada en la encuesta (p. ej., problemas percibidos de organización, accesibilidad,

competencia profesional, trato, etc.), expresándolo como porcentaje de problemas, según la fórmula:

$$\frac{\Sigma \text{ respuestas que expresan existencia del problema}}{\Sigma \text{ preguntas respondidas}} \times 100$$

En esta aproximación, los ítems individuales pueden servir como indicadores específicos de los diferentes aspectos del concepto, dimensión o variable más compleja que resume el indicador compuesto^{10,11}. El principal y quizás único inconveniente que pueden tener los indicadores tipo porcentaje simple es que su resultado pueda ser dominado por el subindicador que sea más frecuentemente evaluado (o la pregunta más frecuentemente contestada, en el caso de las encuestas de calidad percibida); no obstante, este sesgo potencial podría ser, hasta cierto punto, deseable, ya que reflejaría más fielmente la realidad si lo más frecuente está más representado.

Indicadores compuestos con reconversión de la escala de medición de los subindicadores

Uno de los problemas que puede haber al agregar varios subindicadores en uno solo para compararlos es que se dé alguna de las siguientes circunstancias: a) los subindicadores no tienen la misma unidad de medida; b) los subindicadores tienen diferente rango posible de valores máximo y mínimo, y c) los subindicadores tienen diferentes niveles de precisión y dispersión de los datos, en función de haber sido medidos sobre tamaños de muestra (o características de ésta) también diferentes. Todos estos problemas van a restar fiabilidad al indicador compuesto, hasta el punto de poder hacerlo inconsistente y poco útil para comparaciones. Cuando se detecta o se piensa que puede ocurrir alguna de estas circunstancias, o simplemente como medida preventiva por si llegara a ocurrir, se realiza una reconversión de las escalas de medida de los subindicadores, normalizándolos y haciéndolos más comparables.

El método más común de reconversión es dividir los valores observados en cada subindicador por alguna medida que represente la dispersión de los datos, como puede ser el rango o, más frecuentemente, la desviación estándar del valor

Tabla 4. Indicador compuesto tipo porcentaje simple: atención de calidad en pacientes con infarto de miocardio: (casos evaluados: 15)¹⁸

| Subindicadores | Numerador (cumplimientos) | Denominador (casos en los que se ha evaluado) | Cumplimiento |
|---|------------------------------|--|--------------|
| Aspirina dentro de las primeras 24 h | 10 | 12 | 0,83 |
| Bloqueadores beta dentro de las primeras 24 h | 4 | 6 | 0,66 |
| Aspirina al alta | 4 | 5 | 0,80 |
| Bloqueadores beta al alta | 2 | 4 | 0,50 |
| Consejo antitabaco | 1 | 3 | 0,33 |
| Trombolíticos dentro de los primeros 30 min | 7 | 10 | 0,70 |
| Indicador compuesto (calidad de la atención) | 28 | 40 | 0,7 |

Adaptada de Courtney et al¹⁸.

promedio (error estándar de la medición). Por ejemplo, para la construcción de los indicadores compuestos que resumen los datos del Consumer Assessment of Health Plan Study (CAHPS) y Health Plan Employer Data Information System (HEDIS)⁵ (tabla 2), los ítems del CAHPS tenían escalas de respuesta diferentes (0-10, 1-4, 0-1, 1-3) y los indicadores del HEDIS, diferencias también apreciables en su varianza; la solución adoptada fue normalizar todos los subindicadores, antes de su agrupación, dividiéndolos por la desviación estándar de su respectiva media.

En las últimas ediciones del listado Top 20 de los hospitales españoles¹², también se introduce en cierta forma este concepto al normalizar el *ranking* de los hospitales en cada uno de los 6 indicadores por la distancia del valor observado a los valores de los otros hospitales de la lista; estos valores normalizados son los que se suman para obtener el indicador (*ranking*) compuesto final.

En ocasiones se realiza un ajuste con relación a un estándar de referencia. Por ejemplo, en el caso de los indicadores de alto nivel del National Health Service (NHS) británico¹³, el indicador compuesto es la media de las tasas de los subindicadores estandarizados con relación a la tasa promedio de Inglaterra; así, el indicador “cobertura de detección precoz de cáncer en la mujer” que incluye los subindicadores referentes a las coberturas para cáncer de mama (tasa 1) y de cáncer de cérvix (tasa 2) se calcula cómo:

$$\left[\left(\frac{100}{\text{tasa 1 en Inglaterra}} \times \text{tasa 1 en área} \right) + \left(\frac{100}{\text{tasa 2 en Inglaterra}} \times \text{tasa 2 en área} \right) \right] / 2$$

La tabla 5 contiene ejemplos de indicadores compuestos de este tipo utilizados en el NHS británico. En todos ellos se realiza el cálculo utilizando este tipo de estandarización.

Indicadores compuestos con ponderación de los subindicadores

La cuestión de la normalización de las escalas de los subindicadores puede considerarse, aunque sea simplificando mucho el argumento, una cuestión de razonamiento estadístico; es decir, valorar si existe alguna de las circunstancias que lo harían necesario (descritas en el apartado anterior) y encontrar la fórmula de reconversión de los datos que mejor se aadecue al indicador compuesto de que se trate. Sin embargo, plantearse la necesidad de ponderación y decidir qué método utilizar para ponderar raramente va a estar exento de una discusión que va mucho más allá de la fórmula matemática que se emplee. Para empezar, al igual que no hay ninguna forma inequívoca de decidir qué y cuántos subindicadores incluir en un indicador compuesto, a la hora de ponderarlos tampoco existe un procedimiento universalmente aceptado. Todos ellos implican algún tipo de juicio de valor. Vamos a revisar brevemente los más comúnmente empleados.

La utilización de la opinión de expertos o de la población en general tiene un cierto predicamento, sobre la base de su, al menos aparente, legitimidad. La opinión de expertos se ha utilizado, por ejemplo, para ponderar las 6 categorías de subindicadores en el caso de Canadá³ (tabla 1); y el resultado de encuestas a la población se ha utilizado en el caso de los subindicadores considerados por el King's Fund, y de los del indicador compuesto de la OMS⁶, aunque en este último caso la encuesta se dirigió a individuos supuestamente informados y conocedores del tema (la mayoría de ellos emplea-

Tabla 5. Ejemplos de indicadores compuestos: indicadores de alto nivel del National Health Service británico

| | |
|---|--|
| Cobertura de detección precoz de cáncer en la mujer | Promedio de tasas estandarizadas de: Cobertura de detección precoz de cáncer de mama Cobertura de detección precoz de cáncer de cérvix |
| Manejo de patologías de hospitalización evitable | Tasa de admisión estandarizada por: Infecciones renales y del tracto urinario Insuficiencia cardíaca Infección ORL grave |
| Manejo de patologías crónicas | Tasa de admisión estandarizada por: Asma Diabetes Epilepsia |
| Atención de salud mental en atención primaria | Consumo de benzodiacepinas Razón antidepresivos/benzodiacepinas |
| Prescripción coste-efectiva | Coste/prescripción de asociaciones Coste/prescripción de fármacos modificados Coste/prescripción de fármacos de valor clínico limitado Coste/DDD de corticoides inhalados Razón antidepresivos/benzodiacepinas |
| Alta hospitalaria | Tasa de altas < 56 días desde ingreso por accidente cerebrovascular Tasa de altas < 56 días desde ingreso por fractura de cuello de fémur |

DDD: dosis diaria definida.

Tabla 6. Indicadores de calidad en el nuevo contrato de atención primaria en el National Health Service. Puntuación otorgada a cada grupo para su resumen como indicador compuesto

| Grupo de indicadores | Puntos otorgados (máximo) |
|---------------------------------|---------------------------|
| Clínicos (incluye 10 subgrupos) | 550 |
| Organizacionales | 184 |
| Experiencia del paciente | 100 |
| Servicios adicionales | 36 |
| Atención integral | 100 |
| Calidad del centro | 30 |
| Acceso | 50 |
| Total (compuesto total) | 1.050 |

dos de la propia OMS). Un proceso semejante de consulta a supuestos conocedores del tema se realizó por el extinto INSALUD para establecer la ponderación, que llamaron “valor técnico”, de las coberturas de los diversos servicios de la Cartera de Servicios de Atención Primaria, que eran sumadas junto con el cumplimiento de las Normas Técnicas Mínimas, para calcular el “producto” de los centros de salud^{9,14}. También dentro de este mismo grupo de indicadores compuestos con ponderación por la opinión de expertos merece la pena destacar la reciente iniciativa del NHS británico en su nuevo contrato con el sistema de salud¹⁵. En esta iniciativa se resumen en puntuaciones primero parciales, por grupos de indicadores, y luego en una sola puntuación global (que se traduce en incentivo económico: cada punto vale 75 libras esterlinas en 2004/2005 y se aumentará a 120 libras esterlinas en 2005/2006) los indicadores de calidad de la atención primaria. Los grupos de indicadores considerados en esta iniciativa son los que figuran en la tabla 6. Dentro de estos grupos puede haber subgrupos (como en el caso de los indicadores clínicos, con 10 subgrupos correspondientes a los 10 problemas de salud o tipo de pacientes que se considera), y hasta 4 formas diferentes de calcular la puntuación que resume el resultado de lo que miden los subindicadores. Por ejemplo, para la mayoría de los indicadores clínicos el nivel de cumplimiento de los subindicadores se traduce en un número de puntos que va desde 0 a un máximo preestablecido y que se corresponden con un mínimo de cumplimiento por debajo del cual no se puntúa y el estándar de cumplimiento del subindicador, que otorga el máximo de puntuación si se alcanza o sobrepasa; niveles intermedios de cumplimiento se traducen en el número proporcional de puntos desde 0 al máximo. Sin embargo, los subindicadores organizacionales, de experiencia del paciente y servicios adicionales son del tipo sí/no, y la puntuación se otorga en el caso de que estén presentes. La tabla 7 ejemplifica estos 2 procedimientos, con el indicador compuesto de tipo clínico para la calidad de la atención al hipertenso y el de la experiencia del paciente.

Otro ejemplo, aunque puede que no esté concebido expresamente como tal, es la puntuación que se otorga, tanto

por criterios como global, en la evaluación con el modelo EFQM (European Foundation for Quality Management). Este modelo sigue la misma filosofía de un indicador compuesto, resumiendo la evaluación en una puntuación con ponderación de sus componentes, establecida según opinión de expertos.

La forma más aséptica de ponderar, pero no por ello menos discutida, es la utilización de técnicas estadísticas (análisis de componentes principales, análisis factorial, regresión múltiple u otros). Es el método utilizado, por ejemplo, para encontrar qué agregar y con qué peso los 23 subindicadores CAHPS y HEDIS⁵, a los que se sometió a un análisis factorial para identificar los 4 indicadores compuestos que finalmente se calculan. Conviene señalar que 5 de los 9 subindicadores del CAHPS son a su vez indicadores compuestos tipo porcentaje simple (tasa de problemas) procedentes de una encuesta de calidad percibida, asumidos con toda normalidad como subindicadores simples (aunque en realidad son compuestos) y válidos.

Aunque todavía no se han contrastado con la práctica empírica en servicios de salud, las 2 formas de ponderación y agregación de subindicadores que probablemente merecerían ser especialmente consideradas son las que podríamos llamar de “distancia a la frontera de eficiencia” y “distancia a valores estándar”. Ambos métodos han sido utilizados en el ámbito socioeconómico, pero aún no se les ha prestado suficiente atención en la gestión de la calidad de los sistemas de salud, a pesar de que implican conceptos muy arraigados en este campo, como son la referencia a valores *benchmark* (“frontera de eficiencia”) y a valores deseables u óptimos de calidad (“distancia a valores estándar”). Un ejemplo de este tipo de ponderación es la utilizada en el Índice de Desarrollo Humano que utiliza Naciones Unidas para comparar países; en este indicador compuesto, los valores de los subindicadores (esperanza de vida, nivel de analfabetismo y escolarización, PIB per cápita) se expresan como decimales que representan su distancia al valor 1, que es el que se adjudica a las metas que se considera que se debe alcanzar en estos subindicadores (esperanza de vida de 85 años, 100% de escolarización y ausencia de analfabetismo, y PIB per cápita de 40.000 dólares).

Tanto la “distancia a la frontera de eficiencia” como la “distancia a valores óptimos” pueden ser especialmente útiles a efectos de comparar centros, áreas, regiones o sistemas, dentro de una dinámica de mejora continua. Debido a este hecho, junto a la fácil comprensión y elevada validez de apariencia de ambos métodos, se están incorporando a experiencias piloto dentro del Programa EMCA de la Región de Murcia.

Atributos deseables de los indicadores compuestos

Como resumen de todo lo expuesto hasta aquí con relación a los indicadores compuestos, y atendiendo a lo que se sabe en este momento, puede subrayarse una serie de atributos o características a valorar y exigir a este tipo de herramientas. Estas características se describen en la tabla 8. Los 2 primeros atributos (correcta selección de subindicadores y la facilidad de interpretación) están, inicialmente,

Tabla 7. Ejemplos de reconversión a una puntuación ponderada en el cálculo de indicadores compuestos en el nuevo contrato de atención primaria en el National Health Service británico

| A. Ejemplo de indicador compuesto clínico: atención al paciente hipertenso | | | | |
|--|--|---|--------------------------|-------------------|
| Subindicadores | % por debajo del cual puntúan como cero | % a partir del cual se obtiene el máximo de puntos | Puntuación máxima | |
| 1. Registro: El centro tiene un registro de pacientes hipertensos | – | – | 9 | |
| 2. Diagnóstico y manejo inicial | | | | |
| Porcentaje de pacientes hipertensos en los que consta situación en cuanto al hábito tabáquico | 25 | 90 | 10 | |
| Porcentaje de pacientes hipertensos con hábito tabáquico a los que se ha ofrecido consejo antitabáquico | 25 | 90 | 10 | |
| 3. Seguimiento y control | | | | |
| Porcentaje de pacientes hipertensos con registro de tensión arterial en los últimos 9 meses | 25 | 90 | 20 | |
| Porcentaje de pacientes hipertensos con el último registro (en los últimos 9 meses) ≤ 150/90 mmHg | 25 | 70 | 56 | |
| Total (compuesto) | | | 105 | |
| B. Indicador compuesto sobre experiencia del paciente | | | | |
| Subindicadores | | | | Puntuación |
| 1. Duración de la consulta (se establece una duración de 8 o 10 min, según tipo de consulta) | | | | 30 |
| 2. Realización anual de una encuesta acreditada | | | | 40 |
| 3. Reacción a los resultados de la encuesta proponiendo los cambios oportunos | | | | 15 |
| 4. Los resultados de la encuesta se discuten con grupos de pacientes o sus representantes, se proponen cambios y hay alguna evidencia de que estos cambios se han implementado | | | | 15 |
| Total (compuesto) | | | | 100 |

Adaptada de: UK Department of Health: Investing in General Practice. The New GMS Contract. London: Department of Health, 2003.

sujetos a una valoración relativamente subjetiva. Sin embargo, el resto de atributos que se describen en la tabla deben ser valorados de una forma más objetiva. La forma en que capturan la variabilidad entre las unidades a comparar con

Tabla 8. Atributos deseables de los indicadores compuestos

- Resumir un conjunto de indicadores/criterios teóricamente relacionados
- Ser fácilmente interpretables (alta validez de apariencia)
- Describir la mayoría de la variabilidad entre las unidades evaluadas (profesional, centro, área, etc.)
- Se han construido teniendo en cuenta de forma explícita y razonada:
 - La necesidad o no de normalización de las escalas de medida
 - La necesidad o no de ponderación
 - La adecuación de los métodos eventualmente utilizados para los cambios de escala, ponderación y agregación de los subindicadores

relación a lo que pretenden medir debe ser sujeta a prueba y validación empírica. Asimismo, deben buscarse y explicitarse las razones para la manipulación estadística (normalización, ponderación, etc.) a que se someten los subindicadores. La ausencia de una explicación clara y razonable sobre la ponderación siempre va a añadir motivos para la discusión y el escepticismo a la ya de por sí siempre discutible ponderación en general. Un ejemplo claro en este sentido es la iniciativa de los indicadores a valorar en el nuevo contrato de atención primaria en el NHS, comentados más arriba, en donde no se explica, y ya ha sido criticado¹⁶, por qué se da el peso que se da a cada indicador y grupo de indicadores.

Es previsible y deseable¹⁷ una utilización cada vez mayor de los indicadores compuestos en nuestro país, a medida que se avanza en las iniciativas de monitorización de calidad en las diversas comunidades autónomas. Sería conveniente, sin embargo, que se hiciese con el máximo rigor metodológico posible, para que no se produzcan mediciones y comparaciones internamente inválidas, engañosas y, en definitiva, perfectamente inútiles e incluso perjudiciales.

Bibliografía

1. Saturno PJ. Qué, cómo y cuándo monitorizar: marco conceptual y guía metodológica. Rev Calidad Asistencial 1998;13:437-43.
2. Hurtado MP, Swift EK, Corrigan JM, editors. Envisioning the National Health Care Report. Washington: National Academy Press, 2001.
3. Smith P. Composite Indicators of Health System Performance. En: Smith P, editor. Measuring up: improving health systems performance in OECD countries. Paris: OECD, 2002.
4. State-of-the art Report on Current Methodologies and Practices for Composite Indicator Development. First Workshop on Composite Indicators of Country Performance. Ispra (Italia), mayo de 2003 [consultado 12/2003]. Disponible en: http://web-farm.jrc.cec.eu.int/uasa/events/oecd_12may03/
5. Zaslavsky AM, Shaul JA, Zaborski LB, Cioffi MJ, Cleary PD. Combining Health Plan Performance Indicators into Simpler Composite Measures. Health Care Financ Rev 2002;23:101-15.
6. The World Health Report 2000. Health systems: improving performance. Geneva: World Health Organization, 2000 [consultado 12/2003]. Disponible en: <http://www.who.int/health-systems-performance/wqr2000.htm>
7. García-Eroles L, Illa C, Arias A, Casas M. Los Top 20, 2000: objetivos, ventajas y limitaciones del método. Rev Calidad Asistencial 2001;16:107-16.
8. Normas Técnicas Mínimas. Servicio de documentación y publicaciones del Instituto Nacional de la Salud. Madrid: INSALUD, 1993.
9. Manual de procedimiento para la evaluación de la Cartera de Servicios. Servicio de documentación y publicaciones del Instituto Nacional de la Salud. Madrid: INSALUD, 1993.
10. Saturno PJ. Métodos de administración de una encuesta. Análisis y control del sesgo de no respuesta. Presentación de resultados. Manual del máster en gestión de la calidad en los servicios de salud. Módulo 6: participación del usuario en los programas de gestión de la calidad. Exploración y medición de necesidades, expectativas y satisfacción. Unidad temática 36. Murcia: Universidad de Murcia, 2001.
11. Saturno PJ, Más A, Sainz A, y grupo de investigación Programa EMCA. Nueva encuesta posthospitalización de INSALUD. Murcia: Universidad de Murcia; 2001.
12. Programa Top20: Elaboración del ranking [consultado 11/07/2004]. Disponible en: <http://www.iasist.com>
13. Quality and Performance in the NHS; High Level Performance Indicators and Clinical Indicators. London: NHS Executive, 1999 [consultado 12/2003]. Disponible en: <http://www.performance.doh.gov.uk/indicat/indicat.htm>
14. La Cartera de Servicios de Atención Primaria: desarrollo estructural y metodológico 1991-1999 (mimeo). Madrid: INSALUD, 1999.
15. UK Department of Health. Investing in General Practice: The New GMS Contract. London: Department of Health, 2003 [consultado 15/03/2004]. Disponible en: <http://www.dh.gov.uk/PolicyAndGuidance/OrganisationPolicy/PrimaryCare>
16. Smith PC, York N. Quality incentives: the case of UK General Practitioners. Health Aff (Millwood) 2004;23:112-8.
17. Meneu R. Top be or not Top be. Rev Calidad Asistencial 2001; 16:83-5.
18. Courtney J, Krumholz H, Wang Y, Turnbull B. Using composite measures for the public reporting of hospital performance data. ISQua Conference. Paris, octubre 2002.