

Revista de

ÓRGANO DE LA SOCIEDAD ESPAÑOLA DE CALIDAD ASISTENCIAL

Calidad Asistencial

Director: C. Humet
Editores: G. Carrasco, E. Ignacio
R. Lledó, S. Lorenzo, J.J. Mira,
S. Peiró, F. Solsona

Junta Directiva de la Sociedad
Española de Calidad Asistencial

Presidente: C. Humet
Secretario: S. Lorenzo
Tesorero: Ll. Nualart
Vocales: M. Dotor, A. Fernández León,
P. Parra, R. González-Llinares

Comité de Revisores de Metodología

J. Aranaz (Alicante)
R. Pastor (Madrid)
JM. Quintana (Bilbao)
J. Serra (Tarragona)

Consejo Editorial

F. Abel (Barcelona)	A. Mira (Barcelona)
C. Abelló (Barcelona)	D. Montserrat (Madrid)
A. Anglada (Barcelona)	O. Moracho (Guipúzcoa)
A. Arcelay (Vitoria)	F. Moreu (Barcelona)
M.A. Asenjo (Barcelona)	A. Net (Barcelona)
M. Bacigalupe (Vitoria)	A. Ortega (Pamplona)
J. Bañeras (Barcelona)	V. Ortún (Barcelona)
L. Bohigas (Barcelona)	A. Otero (Madrid)
F. Borrell (Barcelona)	V. Pacheco (Barcelona)
P. Busturia (Bilbao)	F. Palacio (Tarragona)
Ll. Cabré (Barcelona)	A. Pallarés (Barcelona)
A. Campos (Valencia)	T. De Pedro (Pamplona)
J.M. Carbonell (Tarragona)	P. Pérez (Barcelona)
G. Cisneros (Madrid)	A. Prat (Barcelona)
A. Conesa (Barcelona)	J. Ribas (Barcelona)
A. Corbella (Barcelona)	J. Roca (Granada)
R. Cot (Barcelona)	F. Rodríguez (Málaga)
R. Delgado (Barcelona)	J.M. Rodríguez (Palencia)
A. Esteban (Madrid)	P. Rodríguez (Madrid)
P. Fernández (Madrid)	JM. R. Roldán (Madrid)
A. Febrer (Barcelona)	J. Rovira (Barcelona)
J.M. Garcés (Barcelona)	A. Ruiz (Málaga)
A. García de Jalón (Zaragoza)	E. Sánchez (Bilbao)
A. García Rodríguez (Barcelona)	J.P. Sancho (Barcelona)
G. Gelabert (Barcelona)	M. Santiñà (Barcelona)
A. Gil (Barcelona)	C. Sanz (Barcelona)
A. Giménez (Barcelona)	P. Saturno (Murcia)
R. González (Vitoria)	R.M. Saura (Barcelona)
A. Junoy (Barcelona)	M. Serra (Barcelona)
C. Lacasa (Barcelona)	R. Suñol (Barcelona)
C. Latorre (Barcelona)	M. Tejedor (Sevilla)
S. Lorenzo (Madrid)	P. Torre (Barcelona)
J.C. March (Granada)	A. Trilla (Barcelona)
J.M. Martín (Sevilla)	Ll. Tiquell (Andorra)
P. Martínez (Zaragoza)	E. Vega (Sevilla)
	F. Vilanova (Barcelona)
	E. Zavala (San Sebastián)

Sumario

Presentación

S5 Carlos Humet, Rosa Suñol

Editorial

S6 Rosa Suñol, Julio Frenk

Entrevista

S8 Entrevista al Profesor Avedis Donabedian. Rosa Suñol

Comentarios y artículos

S10 Estructura, proceso y resultado de la atención sanitaria. Susana Lorenzo

S11 Evaluación de la calidad de la atención médica. Avedis Donabedian

S28 Criterios y métodos de evaluación de la calidad asistencial. Julio Frenk

S29 La Calidad de la Atención Médica. Avedis Donabedian

S39 Calidad y Economía. Lluís Bohigas

S40 Calidad, coste y salud: un modelo integrador. Avedis Donabedian

S53 La influencia de los factores socioeconómicos sobre la calidad asistencial. José J. Mira

S54 Epidemiología de la calidad. Avedis Donabedian

S63 Gestión del conocimiento y calidad asistencial. Enrique Ruelas

S64 La investigación sobre la calidad de la atención médica. Avedis Donabedian

S67 Relevancia de los criterios y estándares para la evaluación y mejora de la calidad.
Victoria Ureña

S68 Criterios y estándares para evaluar y monitorizar la calidad. Avedis Donabedian

S79 Reflexiones sobre calidad asistencial. Rosa Suñol

S80 La calidad de la asistencia. ¿Cómo podría ser evaluada? Avedis Donabedian

S88 La necesaria colaboración entre evaluación de tecnologías y evaluación de la calidad
asistencial. Genis Carrasco

S89 Evaluación de tecnología y calidad. Estudio comparativo de evidencias y ambigüedades.
Avedis Donabedian

S95 El equilibrio entre las preferencias individuales y las necesidades sociales en sanidad.
Beatriz Zurita

S96 Los siete pilares de la calidad. Avedis Donabedian

- S101** El papel de los consumidores en la garantía de la calidad en los servicios de salud: una lección de ética. *Pedro J. Saturno*
- S102** Garantía de Calidad en la Atención de Salud: el papel del consumidor. *Avedis Donabedian*
- S107** ¿Quién es el responsable de la gestión de la calidad asistencial. *M. Rashad Massoud*
- S108** La calidad en la atención sanitaria: ¿de quién es la responsabilidad? (1). *Avedis Donabedian*
- S112** Una reconsideración de las responsabilidades profesionales e institucionales ante la garantía de Calidad Asistencial (2). *Avedis Donabedian*
- S117** La continuidad y el cambio a través de los modelos tradicional e industrial de garantía de la calidad. *Susana Lorenzo*
- S118** Continuidad y cambio en la búsqueda de la calidad. *Avedis Donabedian*
- S126** Reflexiones sobre el futuro de la calidad asistencial. *Rafael Lledó*
- S127** Prioridades para el progreso de la evaluación y monitorización de la calidad de la atención. *Avedis Donabedian*
- S130** La determinación de hacer que las cosas mejoren. *Salvador Peiró*
- S131** Efectividad de la Garantía de Calidad. *Avedis Donabedian*
- S136** La evaluación de la competencia de los médicos. *Julio Frenk*
- S137** Evaluación de la competencia médica. *Avedis Donabedian*
- S140** Reflexión desde la experiencia. *Andoni Arcelay*
- S141** El pasado y el futuro a los ochenta años. *Avedis Donabedian*
- Poema**
- S144** *Avedis Donabedian*

Directora ejecutiva: Mónica López

Editor: Nexus A.G.

C/ Sicilia 364, Entresuelo

08025 (Barcelona)

Tel.: 93 459 34 92 - Fax: 93 459 02 83

Publicidad: Nexus, S.L.

Correspondencia científica: Revista de Calidad

Asistencial. Redacción. C/ Sicilia 364, Entresuelo

08025 (Barcelona)

Imprime: JNP, S.L. Tel.: 93 459 06 64

ISSN: 1134-282X

Déposito Legal: B-12.571/94

Publicación autorizada como Soporte Válido

Periodicidad: 8 números al año



Publicación adherida a la APP, no
sujeta a control obligatorio de difusión

Revista recogida en: **Índice Médico Español**
EMBASE/Excerpta Medica
IBECs

Profesor Avedis Donabedian

Beirut (Líbano 1919-2000)

El Dr. Donabedian fue Profesor Emérito de Salud Pública por la Universidad de Michigan. Su dilatada labor en la sistematización de conocimientos dentro de las ciencias de la salud, especialmente centrada en la mejora de la calidad asistencial, ha sido reconocida mundialmente y merecedora de múltiples galardones.

Cuando está a punto de cumplirse un año de la pérdida de Avedis Donabedian, la Sociedad Española de Calidad Asistencial y la Fundación Avedis Donabedian presentan este suplemento especial en homenaje a la figura del ilustre profesor.

Esta iniciativa, que partió de la directora de la citada fundación, fue compartida con entusiasmo por nuestra sociedad científica, por los editores de la revista y por distinguidas personalidades del mundo de la calidad.

La consecución de proyecto ha requerido más de 500 correos electrónicos y faxes, más de 100 llamadas telefónicas -algunas de ellas intercontinentales- e incontables horas de arduo trabajo. Pero sin duda, el mayor mérito es el de las personas cuyo esfuerzo y dedicación ha hecho posible este complejo proyecto editorial. A todos ellos queremos expresarles nuestro profundo agradecimiento, haciendo mención especial de aquellos que han tenido un papel clave en su realización. A los profesionales de la Fundación Avedis Donabedian y en especial a Pilar Ivern cuya labor de gestión y coordinación pre-editorial han sido inmejorables. A los traductores de las versiones originales que se han enfrentado al inglés rico y culto de Avedis cuya infinidad de matices literarios son un reto para los mejores profesionales de la traducción. A los miembros de la Junta de la Sociedad Española de Calidad Asistencial con mención especial para Susana Lorenzo por su excelente trabajo y su papel dinamizador del proceso editorial. A los miembros del consejo editorial de la revista, en especial a Ángela Pallarés que ha realizado, con inigualable eficiencia, la difícil tarea de contacto internacional con los editores de otras revistas para conseguir los permisos necesarios. A los editores de nuestra revista y muy especialmente a Genís Carrasco que ha llevado la pesada carga de la gestión editorial dedicando muchas horas de trabajo en la supervisión y adaptación del formato de los manuscritos. Al equipo de Nexus Ediciones, con el mayor agradecimiento para nuestra gran periodista científica Mónica López. Ella ha conseguido, con su habitual profesionalidad y eficacia, que lo que parecía imposible sea hoy una espléndida realidad.

Sin duda esta publicación, como toda obra humana, podría ser mejorada en la mayoría de sus aspectos. En la propia selección de artículos que, aunque consensuada con aquellos que conocieron personalmente al profesor, necesariamente tuvo que dejar fuera textos de gran valor metodológico y conceptual. En el orden de presentación de los artículos, que generó un arduo debate sobre si debía ser cronológico o temático, acabando en una decisión salomónica. Y finalmente en el aspecto formal de los comentarios, que preceden a cada trabajo y al propio artículo en sí, cuyo formato hubo que adaptar a las condiciones expresadas por los editores de las publicaciones originales, que amablemente cedieron los derechos de reproducción a nuestra revista sumándose al homenaje al insigne profesor.

Este suplemento compendia por primera vez y en español trabajos muy difíciles de conseguir cuyo valor intrínseco constituye una gran aportación a la literatura científica. Creemos que será un elemento esencial de consulta para todos los que estamos embarcados en esta gran aventura de la calidad que capitaneó el profesor Donabedian.

El lector disfrutará leyendo y releendo estos textos cuyo rigor, lucidez y vigencia transmiten el alma del gran ideólogo de la calidad asistencial. Éste es el mejor homenaje que podemos rendirle.

Carlos Humet

Presidente de la Sociedad Española de Calidad Asistencial

Rosa Suñol

Directora de la Fundación Avedis Donabedian

Este número monográfico, coordinado conjuntamente por la Fundación Avedis Donabedian y el consejo editorial de la revista Calidad Asistencial contiene algunos de los trabajos más significativos del Prof. Avedis Donabedian, recientemente fallecido y a la vez es un homenaje a su memoria.

Por primera vez se publica una selección en español de sus principales trabajos y si bien no están todos los que son (algunos porque su extensión hacía imposible reproducirlos y otros porque forman parte de libros y pierden su coherencia si no se leen en conjunto), si son todos los que están y permiten una buena perspectiva de la magnitud de su obra.

Avedis Donabedian (1919-2000) fue todo un personaje. Sus trabajos en el campo de la Salud Pública son una aportación gigantesca y en el ámbito de la calidad representaron el punto de inflexión en el desarrollo de una ciencia incipiente. Pero el Prof. Donabedian fue muchas cosas más: mentor y guía para muchos, riguroso analista intelectual, enamorado de la lengua y las formas de expresión (hablaba y leía más de 7 idiomas), poeta y sobretodo un hombre de extraordinaria calidad humana, preocupado por los problemas de sus amigos y del mundo y que tendió puentes entre tendencias y culturas diversas.

De origen armenio -sus padres huyeron del genocidio a través del desierto- nació en Beirut (Líbano) y se estableció en Jerusalén. Una nueva guerra le llevó a Estados Unidos, primero a Harvard y posteriormente a Ann Arbor (Michigan) donde desarrolló toda su carrera en la Escuela de Salud Pública.

Su aportación fue básicamente conceptual. A partir del análisis de los trabajos publicados estableció la famosa clasificación de los métodos de calidad (estructura-proceso-resultado), la definición de calidad, la relación entre métodos de proceso y resultado, la sistematización de los criterios, reflexiones muy importantes sobre la responsabilidad en la mejora de calidad, los diferentes enfoques para gestionarla etc. Además terció en dos de las polémicas más importantes del siglo (el debate entre métodos de proceso y resultados y la polémica entre garantía de calidad y calidad total) con artículos cruciales que sintetizaban las posturas y permitían avanzar.

Una de sus aportaciones más constantes es la reflexión del componente ético que define las relaciones en el ámbito de la calidad. Debatía no solo la responsabilidad ética con los pacientes, sino la que afecta a los profesionales individuales, a los colegios profesionales, a los gestores y a la administración, aportando en todos los casos una visión profunda y global.

Si tuviéramos que destacar alguna de sus características científicas creo que deberíamos mencionar especialmente su rigor intelectual y su pasión por la expresión exacta de los conceptos. En muchas ocasiones, intentando seguir una explicación de algún ponente poco riguroso se le veía hacer un esfuerzo físico, casi doloroso, para entender lo que quería decir, hasta que una vez convencido que aquello no tenía sentido, se encogía de hombros como diciendo «no vale la pena». Con sus alumnos era generoso, siempre dispuesto a dedicarles tiempo y a debatir sus trabajos, pero exigente en la discusión. No aceptaba que los que consideraba sus discípulos pensaran sin orden lógico.

Su pasión por la expresión, le llevó a utilizar un idioma rico y complejo, nada fácil de traducir y que esperamos haber reflejado en esta publicación. Creemos que esto se explica también por su pasión por la literatura, especialmente la española. Fue un impenitente lector de nuestro Siglo de Oro, de Santa Teresa, San Juan de la Cruz y de los poetas medievales que leía a diario y discutía con pasión. Escribió también poesía, al principio y al final de su vida y por ello hemos querido aportar también en este número una de sus últimos poemas.

Todo el que trabaja a nivel conceptual tiene como patrimonio único las ideas que aporta. Y no es fácil el que se reconozca la paternidad de las ideas en un mundo en el que las ciencias aplicadas y los resultados tangibles se valoran mucho. Avedis luchó para que se vieran reconocidas sus aportaciones y no fue en su época inicial un hombre fácil. Posteriormente, la edad, la experiencia y su calidad humana hicieron que fueran mucho más importantes para él las personas que las ideas, los lazos que los debates. Sus discípulos encontraron en su casa no sólo cuidadoso soporte científico sino un trato de amigos, con largas sobremesas en las que se hablaba de lo divino y de lo humano con gentes de todos los países y opiniones y una magnífica acogida de su esposa Dorothy.

Sus publicaciones abarcan más de 20 libros y 100 artículos, entre ellos su obra clave «Explorations in quality assessment and monitoring» en tres volúmenes y más de artículos traducidos a muchos idiomas. Recibió numerosos reconocimientos. Era Nathan Sinai Distinguished Professor Emeritus de Salud Pública, miembro del Instituto de Medicina de la Academia Nacional de Ciencias de Estados Unidos, miembro honorario del Royal College of General Practitioners en el Reino Unido y de la Academia Nacional de Medicina de México además de estar en poder de la Medalla Sedgwick por Servicios distingui-

dos en Salud Pública, la condecoración más alta que otorga la American Public Health Association. Además existen con el afán de continuar su labor la Fundación Avedis Donabedian en Barcelona y Argentina y las bibliotecas Donabedian en Ancona (Italia) y Jerusalén.

Avedis fue además, un hombre de profunda fe que manifestaba de forma austera y con un amor intenso que transmitía en cada uno de sus actos. Contaba con un humor inteligente y socarrón que hacía que las relaciones con los demás fueran cálidas y difícilmente olvidables. Así, al final de sus días, cuando todos sabíamos de su gravedad recibía llamadas de sus amigos de todo el mundo y procuraba consolar y transmitir la paz que llegó a alcanzar.

Esta publicación pone al alcance de los lectores de habla hispana una importante colección de sus artículos que permiten comprender la dimensión del autor en sus diferentes facetas,

útil tanto para los que se inician como para los interesados en profundizar en los diferentes conceptos. Esta edición ha sido posible gracias a la coordinación compartida con el Dr. Genís Carrasco y al soporte de los Dres. Beatriz Zurita y Enrique Ruelas de México, Susana Lorenzo y Pedro Saturno de España y Dante Graña de Argentina. Estamos seguros que será una referencia para todos.

Rosa Suñol

*Directora de la Fundación Avedis Donabedian
y de la Cátedra Donabedian de la Facultad
de Medicina de la Universidad Autónoma de Barcelona*

Julio Frenk

Secretario de Salud. México

Nota

Este suplemento ha sido financiado por MSD, que en todo momento ha respetado su elaboración. Los contenidos del mismo son responsabilidad exclusiva de los editores invitados, la Fundación Avedis Donabedian y la *Revista de Calidad Asistencial*.

Esta entrevista al Profesor Avedis Donabedian fue realizada en Ann Arbor, Michigan (EE.UU), por Rosa Suñol, Directora de la Fundación Avedis Donabedian

¿Cómo se inició su interés por la calidad asistencial?

Yo estudié Medicina en la Universidad Americana de Beirut (Líbano), la ciudad en que nací. Después ejercí la profesión durante algunos años en Jerusalén. En aquellos años la situación en Oriente Medio era muy difícil para las personas que como yo eran de origen armenio, así es que me trasladé a vivir a los Estados Unidos. Una vez allí, empecé a especializarme en Salud Pública en la Escuela de Harvard. En aquellos tiempos yo era un estudiante relativamente mayor, de unos 35 años. Durante mis estudios escribí un trabajo sobre calidad asistencial para mi profesor Franz Goldmann. Después de mi graduación entré a formar parte de un equipo de investigación, liderado por Leonard Rosenfeld, que trabajaba sobre la demanda de la atención sanitaria en el área de Boston.

Uno de los estudios dirigidos por el Dr. Rosenfeld comparaba la calidad de la atención en diferentes hospitales. A pesar de que yo no estaba directamente involucrado en este estudio, tuve la oportunidad de observar y aprender lo que estaban haciendo. Este fue el motivo por el que algunos años después, al principio de mi incorporación a la Universidad de Michigan, el Dr. Rosenfeld me sugirió a mí como la persona que podría escribir un artículo de revisión sobre la literatura de evaluación de calidad. Este artículo era parte de un grupo de trabajos encargados a varios autores por una agencia del Gobierno.

En aquella época, la bibliografía sobre temas de calidad era bastante limitada, lo que me permitió leer casi todos los artículos publicados y hacer una ficha de cada uno de ellos. Al intentar clasificar los artículos me di cuenta de que existía una estructura que permitía organizar todo el "saber" que había hasta el momento en los temas de calidad. Sin embargo yo estaba bastante inseguro de que el artículo fuera un éxito por lo que lo mandé con ciertas disculpas. Pero la persona que lo recibió dijo: "chico, has escrito un clásico". Como ustedes saben, el artículo "Evaluating the quality of medical care" (*Milbank Mem Fund Q* 1966;44:166-206) se publicó en el año 1966.

En un principio no pareció que hubiera acertado. El artículo fue ignorado durante 10 años, pero más tarde, ha sido uno de los más citados en nuestro campo y en el área de Salud Pública.

Al releerlo, siempre me ha sorprendido que estuvieran apuntados casi todos los temas que posteriormente he ido desarrollando en mi vida profesional; quizás queden fuera los temas de relación calidad-coste y las medidas de evaluación de resultados.

¿Cuáles cree usted que han sido, posteriormente, las contribuciones teóricas más significativas en el área de evaluación de la calidad en los últimos años?

Creo que se pueden señalar diversos aspectos. En primer lugar, a nivel metodológico, las técnicas que mejoran la priorización y la formulación de criterios: la técnica nominal de grupo (Delbecq), el método Delphi, el desarrollo de la metodología de consenso, etc., han representado un avance muy significativo en metodología de la evaluación de la calidad.

También hay que comentar que el desarrollo del sistema de análisis de decisiones clínicas ha contribuido de forma definitiva a identificar de forma más definitiva las características óptimas de la atención técnica pero que puedan ser incorporadas a los criterios y standards de garantía de calidad.

A nivel de otros ámbitos, es importante destacar el desarrollo de lo desarrollado de los análisis coste-beneficio, de las medidas del estado de salud, sobre todo de los métodos que permiten evaluar las preferencias de los pacientes ante determinados resultados clínicos que han constituido un factor fundamental en la actual prioridad hacia la valoración de los resultados. Otro aspecto importante ha sido, sin duda, el crecimiento de los métodos de estandarización y agrupación del *case-mix*. Estos métodos, además de ser útiles para la gestión y financiación de centros, son útiles para evaluar la calidad ya que hacen posible comparaciones más precisas y válidas de los resultados de la asistencia en diferentes centros.

¿Qué aspectos cree usted que son importantes desarrollar en el futuro?

Hay muchos. En primer lugar, yo recomendaría profundizar en los temas de efectividad de los programas de calidad y su organización, comparando los resultados de los diferentes tipos de programas de mejora de la calidad y su incidencia en la práctica clínica. En segundo lugar, se han de mejorar los conocimientos sobre el cambio organizativo y los métodos para incidir en el comportamiento de los profesionales.

Otro aspecto que me parece importantísimo es generalizar una visión más global de los temas de calidad que esté básicamente enfocada a los temas de preferencia del paciente, pero que integre la valoración de la sociedad frente a los valores personales del paciente, armonice los temas de proceso y resultado, integre los diferentes ámbitos de atención (asistencia

primaria, hospitalaria, centros socio-sanitarios) y se prolongue de forma longitudinal durante las diferentes fases de la vida del paciente.

También faltará hacer compatible el aspecto de participación de los profesionales con un enfoque más centrado en la gestión y la valoración de calidad por organismos externos, con la valoración interna de las instituciones.

Mi opinión es que es necesario un importante debate y un gran esfuerzo de investigación y reflexión para llegar a un acuerdo de quién hace qué en el ámbito de la calidad, evitando duplicaciones y una presión innecesaria sobre los clínicos. Como se puede ver hay trabajo para todos si nos decidimos a avanzar en esta dirección.

Para acabar, ¿qué siente ante las demandas que le llegan de todo el mundo para que vaya a dar clases?

Generalmente mi reacción es positiva. Me gusta explicar personalmente los conceptos a los alumnos y también me gusta viajar. En el Reino Unido y en España me siento como en casa. Con España tengo, además, un lazo afectivo e intelectual: creo que su historia y literatura son únicas y, ahora, con la creación de la Fundación que lleva mi nombre con sede en Barcelona, me siento aún más identificado con este país. Últimamente, no obstante, he viajado mucho, gracias a los amigos europeos y, de momento, me he propuesto pasar una temporada más tranquila en mi casa para tener tiempo de editar y publicar mis lecciones.

Estructura, proceso y resultado de la atención sanitaria

Donabedian A. Evaluating the quality of medical care. *Milbank Mem Fund Q* 1966;44:166-203

Este artículo constituye uno de los hitos en la historia de la evaluación de la calidad de la asistencia sanitaria, introduciendo una de las primeras definiciones de *proceso asistencial*, como un continuo que desagrega en: estructura, proceso y resultado¹. A lo largo del trabajo el Profesor Donabedian describe de forma detallada y ampliamente documentada los métodos utilizados para estimar el grado de calidad de la asistencia sanitaria, proponiendo algunas recomendaciones para el futuro.

El autor se centra casi exclusivamente en la interacción médico paciente para la evaluación de procesos de asistencia sanitaria, a través del análisis pormenorizado de los trabajos de Sheps, Peterson y Lerder, en los que encuentra los fines de la evaluación, los problemas de definición, los criterios y patrones, los sistemas de medida, la fiabilidad de las evaluaciones cualitativas y los índices de calidad, así como la necesidad de definición del episodio de atención del paciente y de la utilización de métodos válidos y fiables para la medida y evaluación de la asistencia sanitaria, basados en definiciones y criterios objetivos.

En cuanto a las definiciones de calidad el Profesor Donabedian establece cómo dependerá el método y el enfoque para su evaluación de las dimensiones y los criterios que se elijan.

¿Qué se debe evaluar? El autor analiza las ventajas de la utilización de resultados como indicadores de calidad de los procesos, al aportar datos concretos, planteando el problema de la fiabilidad de los resultados. En cuanto a la estructura, considera que aunque puede abarcar procesos administrativos de distinta índole, la evaluación así entendida se basaría en la adecuación de las instalaciones y equipos, la idoneidad del personal médico y su organización, la estructura y organización administrativa, etc. El Prof. Donabedian procura aclarar que este análisis de proceso y resultados no significa una separación entre medios y fines, sino una "cadena ininterrumpida de medios antecedentes, seguida de fines intermedios los que, a su vez, son medios para otros fines"; estableciendo así una de las primeras definiciones de proceso asistencial como un continuo.

Considera importante la selección uniforme de aspectos significativos de la atención: cómo se cumplen las tareas en ciertas situaciones clínicas; distinguiendo entre la evaluación de la atención realmente provista y la capacidad de proporcionarla, planteando ya entonces el problema de la variabilidad de la práctica médica y su medida.

En cuanto a las fuentes de información, analiza como el enfoque adoptado determina el método para recoger la información requerida. Sin embargo el autor considera que las cuatro fuentes de información más importantes son: 1. las *historias clínicas* señalando las limitaciones propias de la información que contiene, y que entonces prácticamente solo se utilizaban

para evaluar la atención hospitalaria; 2. *la observación directa por un colega cualificado*; 3. *el enfoque sociométrico*: cuando los médicos buscan atención para ellos y sus familias expresan con sentido crítico opiniones válidas sobre la capacidad de sus colegas de brindar atención de alta calidad; y 4. *el método autorreferencial*: la evaluación que los propios profesionales realizan para juzgar la eficiencia de las organizaciones en las que trabajan.

En cuanto a los patrones de medida, el autor establece dos tipos en función del fin de la evaluación: empíricos y normativos. Considera que cuanto más generales sean los patrones, más dependerán de la interpretación del evaluador. Así mismo analiza los problemas que presentan las distintas escalas de medición aunque, en opinión del Profesor Donabedian, el principal radica en la forma de ponderación de los distintos elementos en el total.

Analiza igualmente validez y fiabilidad: el principal mecanismo para lograr un mayor grado de fiabilidad es la especificación detallada de criterios, patrones y procedimientos empleados para evaluar la atención. Analizando también los problemas derivados de los sesgos de observación. Aunque el Profesor Donabedian considera indudable la relación entre la estructura y el proceso, la complejidad y ambigüedad de las relaciones depende de la gran variedad de factores que intervienen.

A lo largo de todo el trabajo, el autor analiza los diversos enfoques y métodos empleados para evaluar la calidad de la atención médica, señalando una serie de aspectos que merecen consideración, por los problemas que se derivan de ellos. Le parece que se dedica un gran esfuerzo a la formulación de criterios y patrones que presuntamente den estabilidad y uniformidad a cualquier juicio emitido sobre la calidad, y sin embargo, este presunto efecto no se ha demostrado empíricamente.

Aunque el Profesor Donabedian en este artículo se centra casi exclusivamente en la interacción médico paciente en la evaluación de la asistencia sanitaria, plantea la necesidad de búsqueda de formas fáciles de medir un fenómeno tan complejo como la atención médica lo cual podría ser un sueño imposible, aún sin resolver completamente casi cuatro décadas después.

Bibliografía

1. Donabedian A. Evaluating the quality of medical care. *The Milbank Memorial Fund Quarterly* 1966;44:166-203.

Susana Lorenzo
Fundación Hospital Alcorcón

Evaluación de la calidad de la atención médica*

Avedis Donabedian

Introducción

El propósito de este trabajo es describir y evaluar los métodos que se usan actualmente para estimar el grado de calidad de la atención médica, así como proponer algunas orientaciones destinadas a estudios futuros. El trabajo se ocupa de los métodos antes que de los resultados, y de la evaluación de la metodología en general antes que de una crítica detallada de los métodos empleados en determinados estudios.

No se ha querido hacer un examen exhaustivo de la bibliografía sobre el tema. Se han incluido, por cierto, los estudios más importantes; otros se han seleccionado tan sólo como ejemplos ilustrativos. Los omitidos no son por ello menos dignos de mención.

Este trabajo está dedicado casi exclusivamente a la evaluación del proceso de atención médica desde el punto de vista de la interacción entre el médico y el paciente. Por consiguiente, se han excluido los procesos relacionados fundamentalmente con la provisión efectiva de atención médica en el nivel de la comunidad. Tampoco se refiere este análisis a los aspectos administrativos del control de la calidad. Muchos de los estudios que se analizan aquí se originaron en la necesidad imperiosa de evaluar y controlar la calidad de la atención en los programas de atención médica ya establecidos. Sin embargo, estos estudios se examinarán sólo en términos de su contribución a los métodos de evaluación y no en función de sus objetivos sociales más amplios. El autor se ha mantenido, en general, dentro del territorio conocido de la atención que proporcionan los médicos y ha evitado incursionar en otros tipos de atención de la salud. Tampoco se ha considerado el difícil problema de la eficiencia económica como dimensión mensurable de la calidad.

Tres estudios de índole general sobre la evaluación de la calidad han sido sumamente útiles en la preparación de esta reseña. El primero es una obra clásica en la que el autor Mindel Sheps hace un excelente examen de los distintos métodos¹. Un ensayo más reciente de Peterson brinda una valiosa evaluación de este tema². En el informe de Lerner y Riedel se analiza un trabajo reciente sobre la calidad y se plantean varias cuestiones de importancia general³.

Definiciones de la calidad

La evaluación de la calidad debe apoyarse en una definición conceptual y operativa de lo que significa la "calidad de la

atención médica". En este aspecto fundamental se presentan muchos problemas, pues la calidad de la atención es una idea extraordinariamente difícil de definir. Tal vez la definición más conocida sea la de Lee y Jones⁴ en forma de ocho "artículos de fe", algunos enunciados como atributos o propiedades del proceso de la prestación de atención y otros, como metas u objetivos de ese proceso. Estos "artículos" transmiten claramente la impresión de que los criterios de calidad no son nada más que juicios de valor que se aplican a distintos aspectos, propiedades, componentes o alcances de un proceso denominado atención médica. En ese sentido, la definición de calidad puede ser casi cualquier cosa que se quiera que sea, si bien por lo común es un reflejo de los valores y metas vigentes en el sistema de atención médica y en la sociedad más amplia de la que ésta es una parte.

Pocos estudios empíricos analizan en profundidad cuáles son las dimensiones y los valores pertinentes en un momento y en un entorno determinados. Klein *et al.*⁵ encontraron que 24 "funcionarios administrativos" aplicaban, en conjunto, 80 criterios para evaluar la "atención del paciente". Los autores llegaron a la conclusión de que la atención del paciente, al igual que la moral, no se puede considerar como un concepto unitario y "...al parecer es probable que nunca se llegue a un único criterio general para medir la calidad de la atención del paciente".

Las dimensiones y los criterios que se elijan entre los muchísimos posibles para definir la calidad tendrán, por cierto, profunda influencia en los enfoques y métodos que se empleen para evaluar la atención médica.

Modos de enfocar la evaluación: qué es lo que se debe evaluar

El resultado de la atención médica, en términos de recuperación, restauración de las funciones y supervivencia, se ha utilizado frecuentemente como indicador de la calidad de esa atención. Como ejemplos se pueden citar los estudios de la mortalidad perinatal^{6,7}, las tasas de letalidad en la cirugía⁸ y la adaptación social de los pacientes dados de alta de los hospitales psiquiátricos⁹.

El uso de los resultados como criterio para estimar la calidad de la atención médica ofrece innumerables ventajas. Pocas veces se cuestiona la validez de utilizar los resultados

*Donabedian A. Evaluating the quality of medical care. *Milbank Mem Fund Q* 1966;44:166-203

Reproducción autorizada por los propietarios del copyright para la traducción española en homenaje a la figura del profesor Donabedian. Versión española supervisada por Fundación Avedis Donabedian y Revista de Calidad Asistencial

como dimensión de la calidad. Tampoco la estabilidad y validez de los valores de recuperación, restauración y supervivencia plantean ninguna duda en la mayoría de las situaciones y culturas, aunque tal vez no sea así en todas. Por otra parte, los resultados tienden a ser bastante concretos y, como tales, sujetos aparentemente a una medición más precisa.

Sin embargo, algunas consideraciones restringen el uso de los resultados para medir la calidad de la atención. En primer lugar, habría que determinar si el resultado de la atención es, en realidad, la medida que corresponde utilizar, puesto que los resultados reflejan no sólo el poder de la ciencia médica de alcanzar determinados efectos en ciertas condiciones dadas, sino también el grado en que se ha aplicado, en los casos objeto de estudio la, "medicina científica" como se la concibe actualmente. Pero, precisamente, el objetivo puede ser separar estos dos efectos. En algunas ocasiones, un determinado resultado puede no ser pertinente, por ejemplo, cuando se elige la supervivencia como criterio de éxito en una situación que no es mortal, pero que puede conducir a una condición de salud subóptima o invalidante¹⁰.

Aun en las situaciones en que los resultados son pertinentes y en que se ha elegido como criterio el resultado que corresponde, hay que tener en cuenta las limitaciones. En el resultado pueden influir muchos otros factores además de la atención médica, por lo que, si se quiere llegar a conclusiones válidas, habrá que tener cuidado para mantener constantes todos los demás factores distintos de la atención médica. En algunos casos, es preciso que transcurran largos períodos de tiempo, tal vez decenios, antes de que los resultados pertinentes sean evidentes. En esas ocasiones, los resultados no están disponibles cuando se necesitan para la evaluación y los problemas derivados de la necesidad de mantener la comparabilidad se acrecientan notablemente. Además, la tecnología médica no es totalmente efectiva y con frecuencia no se conoce con precisión el índice de éxitos que es dable esperar en una situación determinada. Por esta razón se deben hacer estudios comparativos de los resultados en situaciones controladas.

Si bien algunos resultados son por lo general fáciles de medir sin lugar a error (la muerte, por ejemplo), la medición de otros que no fueron definidos con tanta precisión puede resultar difícil. Entre estos se incluyen las actitudes y satisfacciones de los pacientes, la readaptación social y la discapacidad y rehabilitación física¹¹. Aun la validez aparente que se atribuye por lo común a los resultados, como criterios de éxito o fracaso, no es absoluta. Es una cuestión debatible, por ejemplo, si la prolongación de la vida en determinadas circunstancias es evidencia de buena atención médica. McDermott *et al.* han demostrado que, si bien la corrección de una luxación congénita de cadera en una posición determinada es considerada una prueba de buena medicina en el hombre de raza blanca, puede resultar invalidante para el indio navajo que pasa gran parte de su tiempo sentado en el suelo o sobre la montura¹². Por último, si bien los resultados considerados en conjunto podrían ser indicativos de buena o mala atención, no permiten conocer en qué consisten las deficiencias o ventajas a las que se podría atribuir un resultado determinado ni dónde se encuentran.

Al presentar todas estas limitaciones al uso de los resultados como criterios para evaluar la atención médica no se ha

querido demostrar que los resultados son inadecuados como indicadores de la calidad, sino subrayar que se los debe utilizar con discernimiento. Los resultados siguen siendo en general los indicadores definitivos para dar validez a la eficacia y calidad de la atención médica.

Otra forma de enfocar la evaluación consiste en examinar el proceso mismo de prestación, en lugar de los resultados. Esta posición se justifica porque se supone que lo que interesa no es el poder de la tecnología médica de alcanzar los resultados previstos, sino determinar si se ha aplicado lo que ahora se conoce como "buena" atención médica. Las estimaciones se fundan en consideraciones como el grado de adecuación, integridad y minuciosidad de la información obtenida mediante la historia clínica, el examen físico y las pruebas diagnósticas; la justificación del diagnóstico y el tratamiento; la competencia técnica en la aplicación de los procedimientos diagnósticos y terapéuticos, incluso los quirúrgicos; las pruebas del manejo preventivo del paciente en condiciones de salud y de enfermedad; la coordinación y continuidad de la atención; el grado en que la atención resulta aceptable para el destinatario y muchas otras. Este enfoque requiere que se especifiquen con gran detenimiento las dimensiones, los valores y las normas que se usarán en la evaluación. Las estimaciones de la calidad obtenidas sobre esta base son menos estables y definitivas que las derivadas de la medición de los resultados. No obstante, pueden ser más pertinentes a la cuestión que nos ocupa: establecer si se practica correctamente la medicina.

Podría parecer que este análisis del proceso y los resultados significa una mera separación entre medios y fines. Tal vez sería más acertado pensar en una cadena ininterrumpida de medios antecedentes seguida de fines intermedios los que, a su vez, son medios para otros fines¹³. La salud misma podría ser un medio para alcanzar aun otro objetivo. Varios autores han señalado que esta formulación ofrece un enfoque valioso para la evaluación^{14,15}, la que podría considerarse como la medición de los puntos finales del procedimiento e incluirse bajo el encabezamiento general de "proceso" porque se funda en consideraciones similares con respecto a los valores, criterios y validación.

Una tercera manera de abordar la evaluación es estudiar no el proceso de atención en sí mismo, sino las situaciones en que tiene lugar y los medios y procedimientos de los que es el producto. Esto podría denominarse *grosso modo* evaluación de la estructura, aunque también puede abarcar procesos administrativos de distinta índole, que prestan apoyo y orientación a la provisión de atención. La evaluación así entendida se basa en aspectos tales como la adecuación de las instalaciones y los equipos; la idoneidad del personal médico y su organización; la estructura administrativa y el funcionamiento de programas e instituciones que prestan atención médica; la organización fiscal y varios más^{16,17}. Se da por sentado que si se cuenta con las condiciones y los medios adecuados, la buena atención médica estará asegurada. Este enfoque tiene la ventaja de que se refiere, al menos en parte, a información bastante concreta y accesible. Sin embargo, presenta el inconveniente importante de que con frecuencia la relación entre la estructura y el proceso o entre la estructura y el resultado no está bien establecida.

Fuentes de información y métodos para obtenerla

El enfoque que se adopte para evaluar la calidad determina en gran parte los métodos que se usarán para recoger la información requerida. Puesto que estos abarcan toda la gama de los métodos de las ciencias sociales, no se intentará hacer la descripción de cada uno de ellos. Hay cuatro, sin embargo, que merecen atención especial.

Las *historias clínicas* son los documentos de los que se parte para la mayoría de los estudios sobre el proceso de atención médica. No se debe olvidar, sin embargo, que adolecen de varias limitaciones. Dado que los consultorios de la mayoría de los médicos de la práctica privada no son fácilmente accesibles para el investigador y que los registros que estos llevan son por lo demás someros, las historias clínicas sólo se utilizan para evaluar la atención prestada en los hospitales, incluso en los departamentos de medicina ambulatoria, y en los planes de seguro médico. Tanto Peterson¹⁸ como Clute¹⁹ se han referido a las insuficiencias más comunes que se observan en los registros de la práctica médica general. Clute ha señalado, además, que en la práctica general "...la carencia de registros adecuados no es incompatible con el ejercicio de una medicina de buena y hasta excelente calidad..." Por otra parte, un estudio reciente sobre los consultorios privados de una muestra de médicos miembros de la Sociedad de Medicina Interna de Nueva York²⁰ revela que los resúmenes de los registros llevados en ellos se pueden utilizar para obtener estimaciones reproducibles de la calidad de la atención. Sin embargo, es difícil generalizar sobre este hallazgo, porque se refiere a un grupo especial de médicos cuyos registros son probablemente más adecuados que los del término medio. Además, por una u otra razón, la muestra original recogida para este estudio sufrió una pérdida de 61%.

Suponiendo que se tenga acceso al registro y que sea razonablemente adecuado, queda aún por establecer hasta qué punto es veraz y completo. Lembcke¹⁰ ha objetado que se puedan aceptar sin más las afirmaciones clave incluidas en el registro. No sólo ha cuestionado las declaraciones del médico sobre el paciente y su manejo, sino también la validez de los informes de los servicios de diagnóstico. Para verificar las afirmaciones del médico, se puede buscar en el registro, incluso en las anotaciones de las enfermeras, lo que al parecer sea la prueba más válida sobre el verdadero estado de cosas. La validez de los informes se confirma con el examen por personas competentes de las pruebas (películas, registros gráficos, diapositivas) en las que se basan los informes diagnósticos. Aun en las mejores circunstancias, el error del observador en la apreciación del material suele ser un problema²¹. Pero nada supera la incredulidad que produce el hallazgo de Lembcke en un hospital donde determinó que, sobre la base de las muestras de material de legrado, la verdadera incidencia de la hiperplasia uterina era entre el 5 y el 8%, en lugar de un 60 a 65%, según lo informado por el patólogo del hospital. En todo caso, los resultados de la verificación considerada como parte de la evaluación de la calidad deben ser objeto de detenido análisis. No cabe duda de que los errores en los informes diagnósticos se reflejan particularmente en la calidad de servicio de diagnóstico y en la atención que presta el hospital en general. Sin embargo, el desempeño del médico puede merecer un juicio favorable, prescindiendo de que los datos con que trabaja sean o no válidos.

Ello es así cuando el objeto de interés al emitir ese juicio son los principios que orientan las actividades del médico y no la validez absoluta de esas actividades.

Mucho se ha debatido la cuestión de cuán completas son las historias clínicas, así como la posibilidad de establecer si, al evaluar la calidad de la atención sobre la base del contenido de esa historia, se está atribuyendo un valor a la historia clínica o a la atención prestada. La confusión se origina porque el registro de la historia clínica es de por sí una dimensión de la calidad de la práctica médica, así como el medio de información que permite evaluar la mayoría de las restantes dimensiones. Estos dos aspectos se pueden separar cuando se dispone de una fuente alternativa de información sobre el proceso de la atención como, por ejemplo, la observación directa de la práctica médica^{18,19}. Rosenfeld²² resolvió el problema de separar el registro de la historia clínica y la atención médica prestada examinando las razones por las cuales se le adjudicaba un valor menor a la calidad de la atención en la historia de cada paciente examinada. Demostró que el hecho de que la calidad de la atención se considerase de menor valor se debía en parte a lo que se podía considerar un registro inadecuado (pruebas "presuntas") y en parte a razones que no eran atribuibles al registro (pruebas "reales"). También encontró que la clasificación adjudicada a los hospitales solía ser la misma, alta o baja, con respecto a ambos tipos de errores, lo que indica la existencia de una correlación entre esos errores. Puesto que lo más probable es que los registros de rutina más completos correspondan a las salas, la comparación entre los servicios privados y los de sala de cada hospital en función de los tipos de razones por las que la calidad de estos servicios recibió una clasificación más baja podría haber brindado mayor información sobre esta importante cuestión. Otros investigadores han intentado subsanar las omisiones en las historias clínicas complementando esos informes mediante entrevistas con el médico a cargo del paciente y efectuando las enmiendas correspondientes^{23,25}. Lamentablemente, solo en uno de esos estudios (duración de la ingreso en los hospitales de Michigan) se incluye un informe sobre las diferencias que originó esta medida. En este estudio, "la información médica adicional recogida mediante entrevistas personales con los médicos de cabecera fue lo suficientemente importante en el 12,6% del total de casos estudiados para que se justificara reclasificar la evaluación realizada respecto a la necesidad de ingresar al paciente o a la adecuación de la duración del ingreso, o respecto a ambas"^{3,25}. Cuando la información obtenida en una entrevista se usa para modificar o completar la historia del paciente, es posible que haya que suponer que esta información adicional tiene la misma o mayor validez que la contenida en la historia. Morehead, cuya experiencia en la aplicación de este método es muy amplia, señaló que: "Muchos de los encuestadores que colaboran en el presente estudio habían aplicado el método de entrevistar al médico en trabajos anteriores sin obtener buenos resultados... El encuestador... se encontraba en la incómoda posición de verse obligado a elegir entre aceptar sin cuestionar las aseveraciones de que la atención médica era realmente óptima, o llegar a la conclusión de que esas afirmaciones eran inexactas"²⁶. En un estudio anterior, en el que se indicaba que se recurría al método de las entrevistas para complementar los datos disponibles²⁴, se desechó la

información verbal a menos que la corroborasen las medidas tomadas o pruebas concretas²⁷.

Otra cuestión metodológica que requiere consideración es establecer si corresponde usar como base para la evaluación la totalidad de la historia clínica o resúmenes de ella. La cuestión se plantea porque se supone que los extractos y resúmenes pueden ser preparados por personas menos especializadas a fin de que el experto, cuyos servicios son difíciles de obtener, pueda concentrarse en la verdadera tarea de la evaluación. Sin embargo, la preparación de resúmenes exige discernimiento para decidir qué es pertinente e importante. Por esta razón, sólo se los ha usado como un primer paso para evaluar la calidad en estudios en los que se utilizan criterios muy específicos y detallados¹⁰. Aun así, es poco lo que se sabe acerca del grado de confiabilidad del proceso de extraer información o qué validez tiene cuando se lo compara con una lectura más experimentada del material. El estudio sobre los especialistas en medicina interna de Nueva York al que ya se ha hecho referencia reveló una gran concordancia entre los resúmenes de historias clínicas de consultorios preparados por los médicos y por personas altamente capacitadas que no pertenecían a esa profesión²⁰.

Si bien continúa la controversia acerca de si la historia clínica debe ser utilizada como fuente de información, algunos autores han intentado disminuir su dependencia de los registros médicos eligiendo para la evaluación categorías diagnósticas que estén respaldadas por comunicaciones escritas además de las anotaciones del médico²⁸. Esto explica, en parte, que se recurra con frecuencia a los informes sobre operaciones quirúrgicas como material para los estudios de la calidad de la atención.

Las historias clínicas de los pacientes atendidos en la práctica general son demasiado imprecisas para que se las pueda utilizar como base para la evaluación. La otra alternativa es la *observación directa* de las actividades del médico por un colega calificado^{18,19}. La principal limitación de este método parecería ser los cambios que probablemente se producirían en la actividad habitual del médico que sabe que es objeto de observación. En contra de esto se ha aducido que a menudo el médico no está enterado del verdadero propósito del estudio, que se acostumbra rápidamente a la presencia del observador y que le resulta imposible modificar hábitos arraigados en el ejercicio de su profesión. Aun cuando ocurran cambios, es probable que resulten en una estimación mayor de la calidad y no en lo contrario. A pesar de estas afirmaciones, la medición de los efectos que la observación puede tener en la práctica médica sigue siendo un problema no resuelto.

A quienes han utilizado el método de la observación no se les ha escapado que con él no se obviaba el problema de establecer cuán completa es la información. El médico con frecuencia sabe bastante acerca del paciente por haber tenido contacto con él en ocasiones anteriores. De ahí que sea necesario elegir para la observación casos "nuevos" y situaciones que exijan hacer un reconocimiento a fondo, sea cual fuere la experiencia anterior que se tenga con el paciente. Por otra parte, no todas las actividades del médico para el manejo del paciente son explícitas. Algunos aspectos de la atención, no sujetos a la observación directa, deben excluirse del esquema de evaluación.

Otro problema que se puede presentar es la percepción selectiva del observador. No es lo habitual que el observador primero registre imparcialmente los hechos y sea después el juez de esos mismos hechos. Es muy posible que sus conocimientos y criterios influyan en su manera de percibir los hechos, lo que puede ocasionar una cierta distorsión de lo percibido.

Uno de los métodos indirectos de obtener información consiste en estudiar *conductas y opiniones* que permitan hacer inferencias sobre la calidad. Maloney *et al.* se han referido a un enfoque *sociométrico*, según el cual, cuando los médicos buscan atención para ellos y sus familias expresan con sentido crítico opiniones válidas sobre la capacidad de sus colegas de brindar atención de alta calidad²⁹. Se comprobó que los médicos que eligieron tenían reputación de estar más calificados que otros. Pero tanto la sensibilidad como la especificidad, cuando se usan como criterio estimaciones más estrictas de la calidad de la atención, carecen de validez. Georgopoulos y Mann³¹ emplearon un método, que se podría denominar *autorreferencial*³¹, para evaluar la calidad de la atención en algunos hospitales comunitarios. Este método surgió de estudios anteriores en los que se comprobó que la gente suele ser bastante perspicaz cuando juzga la "eficacia" de las organizaciones en las que trabaja³². Los hospitales fueron clasificados por grado sobre la base de opiniones acerca de la calidad de la atención médica y otras características emitidas por miembros del personal directivo, profesional y técnico de cada hospital, por personas relacionadas con esas instituciones y por gente bien informada de la comunidad. La perspicacia y coherencia de las opiniones permitió que se pudiesen clasificar los hospitales con un grado de confiabilidad aparentemente satisfactorio, a pesar del carácter elogioso de las respuestas que en el 89% de los casos calificaron la calidad de la atención médica en los hospitales de "muy buena", "excelente" o "sobresaliente" y de "regular" en casi ninguno. Los autores suministran abundantes pruebas de un alto grado de interrelación entre distintas opiniones emitidas en forma separada. Pero en pocos casos se respalda la validez de las opiniones recurriendo a criterios verdaderamente independientes para evaluar la calidad de la atención.

Toma de muestras y selección

Cuando se va a efectuar un muestreo, lo primero que hay que hacer es especificar con precisión el universo que será objeto del muestreo, el cual, a su vez, depende de la clase de generalizaciones que se quiera hacer. Por lo común, los estudios sobre la calidad se refieren a uno de estos tres objetivos: 1. la atención que realmente presta una categoría predeterminada de proveedores de atención; 2. la atención que realmente recibe un grupo determinado de personas, y 3. la capacidad de un grupo determinado de proveedores de proporcionar atención. En los dos primeros casos se necesitan muestras representativas tanto de los posibles proveedores o beneficiarios, como de la atención proporcionada o recibida. En el tercero se requiere una muestra representativa de los proveedores, pero no necesariamente de la atención recibida. Otro aspecto aun más importante es que se seleccionen de manera uniforme aspectos significativos de la atención. Tal vez sea necesario estudiar cómo se cumplen las tareas en ciertas situaciones clínicas que originan

mucha tensión y en las que, por consiguiente, se ponen más de manifiesto la competencia o las debilidades latentes para desempeñarlas. Hasta se pueden crear situaciones hipotéticas para evaluar la capacidad de desempeño en ciertos aspectos de la atención³³⁻³⁵. Las distinciones que se han hecho antes, especialmente entre la evaluación de la atención realmente provista y la capacidad de proporcionar atención, son convenientes para evaluar los procedimientos de muestreo utilizados en los principales estudios sobre la calidad. Según estos criterios, ciertos estudios pertenecen a una determinada categoría, mientras que otros reúnen las características de varias categorías de modo tal que la generalización resulta difícil. Por ejemplo, en el primer estudio sobre la calidad de la atención proporcionada a las familias de los miembros del sindicato de conductores de camión, los resultados se refieren sólo al manejo de determinadas categorías de enfermedad que requieren internación con respecto a un grupo de población determinado²⁸. En el segundo trabajo de esta serie se alcanza un grado de generalización algo mayor obteniendo una muestra representativa (con exclusión de las variaciones estacionales) de todas las enfermedades que requieren internación en el mismo grupo de población²⁶. En ninguno de los dos estudios se prevé suministrar información sobre toda la atención prestada por una muestra representativa de médicos.

El grado de homogeneidad del universo que se va a muestrear tiene, por cierto, gran importancia en todo proyecto de muestreo o selección. La pregunta que corresponde formular aquí es en qué medida es uniforme la atención que presta un médico. ¿Existen determinadas categorías diagnósticas, niveles de dificultad o aspectos de la atención en los que un médico se desempeña mejor que en otros? ¿Es posible encontrar realmente en ese médico una "capacidad general de desempeñarse bien al prestar la atención médica"¹⁸ o, por el contrario, la calidad de su desempeño es bastante dispar? También cabría preguntarse si la atención que prestan las distintas subdivisiones de una misma institución es más o menos del mismo nivel ya sea en términos absolutos o en relación con el desempeño de instituciones comparables. Makover, por ejemplo, da por sentada explícitamente la homogeneidad cuando sostiene: "No se hizo ningún intento de relacionar el número de historias clínicas que se iban a estudiar con el número de asociados a los planes de seguro médico. La atención médica prestada a uno u otro individuo es una prueba válida de la calidad y no debe existir ninguna, o sólo una mínima, variación aleatoria que sea afectada por el ajuste según el tamaño de la muestra"²³. Rosenfeld inició su estudio enunciando la hipótesis de que en las instituciones existe una correspondencia en los niveles de atención de las distintas especialidades y para varias categorías de enfermedades²².

No abundan las pruebas empíricas relacionadas con la homogeneidad. Los estudios de Peterson y Clute sobre la práctica general de la medicina^{18,19} mostraron que había un alto grado de correlación entre el desempeño de los médicos en distintos componentes o aspectos de la atención (historia clínica, reconocimiento físico, tratamiento, etc.). Por su parte, Rosenfeld también demostró que no había grandes diferencias en las clasificaciones de la calidad de varios diagnósticos seleccionados en distintos campos de la práctica médica (medicina

general, cirugía y obstetricia-ginecología). Si bien las diferencias entre los hospitales parecieron ser mayores cuando se consideraron los distintos campos de la práctica médica, no fueron suficientes para modificar las posiciones de los tres hospitales estudiados.

En los dos estudios de la atención recibida por las familias de los miembros del sindicato de conductores de camión^{26,28} se citó un porcentaje de atención óptima y subóptima casi idéntico en todas las poblaciones estudiadas. Esa debe haber sido una coincidencia, ya que el porcentaje de atención óptima, en el segundo estudio, varió mucho según la clase de diagnóstico, de 31% en el caso de la medicina a 100% en el de la oftalmología (nueve casos solamente). Si existe esa variabilidad, el "conjunto de diagnósticos" de la muestra de pacientes atendidos debe ser un asunto de gran importancia en la evaluación. En los dos estudios de las familias de los miembros del sindicato citado, se consideró que las diferencias de ese conjunto habían llevado a asignar una menor clasificación a la medicina y una mayor a la obstetricia y ginecología en el segundo estudio en comparación con el primero. El hecho de que el mismo factor pueda producir efectos en dos sentidos contrarios indica la complejidad de las interacciones que debe considerar el investigador. "La explicación más probable de la menor clasificación asignada a la medicina en el estudio actual (que es el segundo) es la naturaleza de los casos analizados". El factor al que puede atribuirse eso es la menor capacidad para manejar la enfermedad "que no encuadró dentro de un patrón bien reconocido". En el caso de la obstetricia y la ginecología, los resultados del segundo estudio "...difieron en un aspecto importante de los del anterior en que se hicieron serias preguntas sobre el tratamiento de un número de pacientes mucho mayor. El estudio anterior se centró sobre todo en cirugía abdominal mayor, en tanto que el grupo seleccionado al azar comprendió pocos de esos casos y tuvo más pacientes con afecciones menores"²⁶. En estudios como éstos, en que la atención recibida por la totalidad o parte de la población es objeto de estudio, las variaciones observadas emanan, en parte, de diferencias de la materia incluida en el diagnóstico y, en parte, de patrones institucionalizados de práctica relacionados con dicha materia. Por ejemplo, los nueve casos de oftalmopatía recibieron óptima atención porque "ese es un campo muy especializado en que los médicos que carecen de la formación pertinente raras veces se arriesgan a practicar procedimientos"²⁶.

El muestreo y la selección influyen en varios factores, además de la generalización y la homogeneidad, y reciben la influencia de los mismos. Las dimensiones particulares de la atención que despiertan interés (tratamiento preventivo o técnica quirúrgica, para citar dos ejemplos bastante distintos) pueden determinar la selección de situaciones de atención médica para fines de evaluación. Las escogidas guardan relación también con la naturaleza de los criterios y patrones usados y el sistema de clasificación y calificación adoptado. Cualquier intento por tomar muestras de situaciones problemáticas, más bien que de diagnósticos u operaciones tradicionales, se puede dificultar mucho por la forma en que se archivan y clasifican las fichas clínicas. Eso es lamentablemente porque un análisis de las operaciones o de los tipos de diagnóstico establecidos da una idea de en qué se fundamentó el diagnóstico o

se hizo la operación. Eso deja inexplorado un segmento complementario de la práctica, a saber, las situaciones en que puede haberse indicado un diagnóstico o tratamiento similar, pero no se emitió ni se administró.

Patrones de medición

La medición depende de la formulación de patrones. Para fines de evaluación de la calidad, los patrones se derivan de dos fuentes.

Los patrones *empíricos* se derivan de la práctica real y se emplean en general para comparar la atención médica prestada en un ámbito con la de otro o con promedios y escalas de valores estadísticos obtenidos a partir de un mayor número de ámbitos similares. El Estudio de las Actividades Profesionales se basa, en parte, en ese método³⁶.

Los patrones empíricos se basan en niveles de atención cuyo logro puede demostrarse y, por esa razón, gozan de un cierto grado de credibilidad y aceptabilidad. Además, sin patrones normativos claros, deben formularse observaciones empíricas en ámbitos seleccionados con esa finalidad. Un interesante ejemplo proviene del trabajo de Furstenberg *et al.*, quienes emplearon patrones de prescripción en clínicos de atención médica y hospitales de atención ambulatoria como norma para juzgar el ejercicio particular³⁷.

Al emplear patrones empíricos hay que tener cierta seguridad de que el material clínico en los distintos ámbitos comparados sea similar. En el Estudio de las Actividades Profesionales esto se tiene en cuenta al notificar los patrones de atención en los hospitales agrupados por tamaño. Sin embargo, la mayor deficiencia está en que la atención puede parecer adecuada en comparación con la que se presta en otros casos y ser inferior a la que puede lograrse con la cabal aplicación de los actuales conocimientos de medicina.

Los patrones *normativos* se derivan, en principio, de las fuentes que fijan legítimamente las normas del conocimiento y la práctica en el sistema de atención médica predominante. En la práctica, son fijados por libros de texto comunes o publicaciones¹⁰, grupos de médicos²⁵, profesionales de máxima idoneidad que sirven de jueces²⁶ o un grupo de investigadores en consulta con profesionales calificados²². Los patrones normativos pueden tener niveles muy altos y representar la "mejor" atención médica que se puede prestar o tener un nivel más modesto que represente una atención "aceptable" o "adecuada". En todo caso, su característica distintiva se centra en que emanan de un conjunto de conocimientos y valores legítimos más bien que de ejemplos específicos de la práctica real. Como tales, su validez depende del grado de concordancia sobre hechos y valores dentro de la profesión o, al menos, entre su personal directivo. Donde las fuentes igualmente legítimas difieren en sus puntos de vista, el juicio relativo a la calidad es también ambiguo.

Se ha puesto en tela de juicio la pertinencia de ciertos patrones normativos, formulados por un grupo, para el campo de práctica de otro. Por ejemplo, Peterson y Barsamian indican que, si bien el examen del líquido espermático del esposo debe preceder a la intervención quirúrgica para tratar el síndrome de Stein-Leventhal, no se practicó ni una sola vez y se suprimió ese requisito de los criterios de evaluación³⁸. Tam-

bién se ha expresado insatisfacción en relación con la aplicación al ejercicio general de patrones y criterios formulados por especialistas que ejercen en el ámbito académico. En los principales estudios sobre la práctica general se ha tenido en cuenta ese asunto. Sin embargo, poco se sabe respecto de las estrategias de la "buena" práctica general y de sus similitudes y diferencias en relación con las de la práctica especializada en el ámbito académico.

Algunos investigadores han empleado ambas clases de patrones, normativos y empíricos, para evaluar la atención. Rosenfeld empleó patrones normativos, pero incluyó en su diseño una comparación entre los hospitales universitarios y los comunitarios. "El empleo del hospital docente como testigo da el elemento de flexibilidad necesario para ajustarse a la siempre cambiante base científica de la práctica de la medicina. Ningún patrón escrito, por bien redactado que esté, sería adecuado en cinco años"²². Lembcke se basó en la experiencia adquirida en los mejores hospitales para formular un factor correctivo que suavice la excesiva rigidez de sus patrones normativos. Ese factor, expresado en términos de un porcentaje aceptable de cumplimiento con el patrón, se formuló para tener en cuenta las situaciones imprevistas en los patrones propiamente dichos. Sin embargo, tiene el efecto de ser más permisivo también en su aspecto real. Eso se debe a que el factor de corrección puede constar, en parte, de desviaciones aceptables con respecto al patrón y, en parte, de otras que podrían ser inaceptables.

Los patrones también se pueden diferenciar por su grado de especificidad y de indicaciones detalladas. En un extremo, el médico evaluador puede recibir instrucciones muy sencillas, por ejemplo, "como una vara que sirva para medir la calidad de la atención prestada, tendrá que determinar si habría tratado a ese paciente de ese forma durante ese período de hospitalización en particular"²⁶. En el otro extremo, se puede formular un "sistema lógico" casi impenetrable en el que se especifiquen todas las reglas sobre decisiones que son aceptables para justificar el diagnóstico y el tratamiento^{38,39}. La mayoría de los casos se sitúan en un punto intermedio.

Los patrones muy precisos y estrictamente detallistas guardan relación con la selección de determinadas clases de diagnóstico para fines de evaluación. Cuando se pretende evaluar una muestra representativa de toda la atención prestada, no se puede dar más que una guía general al evaluador. Por ejemplo, Lembcke, que ha hecho hincapié en la necesidad de formular criterios específicos, ha tenido que establecer una clasificación de diagnóstico de la cirugía de la pelvis igualmente detallada¹⁰. Además de la especificidad del diagnóstico, los patrones estrictamente detallados guardan relación con la selección previa de dimensiones específicas de la atención para efectos de evaluación. Ciertas clases de diagnóstico, como el de intervenciones quirúrgicas, se prestan más a inclusión dentro de este método. Eso es evidente en el intento hecho por Lembcke de ampliar su sistema de verificación al diagnóstico distinto del quirúrgico⁴⁰. Los juicios claros, casi empíricos, de idoneidad pierden su nitidez. Los datos resumidos en cada clasificación de diagnóstico se asemejan más a una descripción de los patrones de manejo, con un número insuficiente de criterios normativos para hacer una evaluación definitiva. En

este caso, se adopta la alternativa de la comparación con una institución con criterio establecido.

Obviamente, cuanto más generales y con menos indicaciones detalladas sean los patrones, más habrá que depender de las interpretaciones y normas de la persona encargada de la evaluación real de la atención. Con mayor especificidad, el grupo investigador puede ejercer, en conjunto, mucho más control de las dimensiones de la atención que se deben subrayar y de lo que constituyen patrones aceptables. Los patrones empleados en situaciones estructuradas y no estructuradas parecen tener mucho en común, según lo indica el grado de concordancia entre la clasificación "intuitiva" y la dirigida en el estudio de Rosenfeld²² y entre la "cualitativa" y "cuantitativa" en el de Peterson y colaboradores¹⁸. En realidad, estos dos últimos grupos fueron tan similares que podrían usarse indistintamente.

Cuando los patrones no son muy específicos y el evaluador debe emitir su propio juicio para poder hacer la evaluación, hay que emplear los servicios de jueces muy expertos y cuidadosos. Lembcke alega que un sistema mucho más preciso y detallado como el suyo no necesita jueces expertos. "Se dice que con un libro de cocina, cualquiera que sepa leer puede cocinar. Lo mismo sucede con la revisión de la atención médica con criterios objetivos, y casi en la misma proporción; cualquiera que sepa suficiente terminología médica para entender las definiciones y criterios puede preparar resúmenes de casos y cuadros para verificación médica. Sin embargo, la aceptación, interpretación y aplicación final de los resultados debe ser la responsabilidad de un médico o un grupo de médicos"⁴¹. El "sistema lógico" formulado por Peterson y Barsamian parece adaptarse bien a la clasificación sistematizada en computadora una vez que los hechos básicos se han recopilado, presuntamente por un especialista en resúmenes de fichas^{38,39}.

Por supuesto, las dimensiones de la atención y los valores empleados para juzgarlas están incorporados en los criterios y patrones usados para su evaluación⁴². Por tanto, esos patrones pueden diferenciarse por selectividad e inclusividad en la elección de las dimensiones que es preciso evaluar. Las dimensiones seleccionadas y el juicio emitido respecto de sus valores constituyen la definición funcional de la calidad en cada estudio.

Como ya se señaló, la selección previa de las dimensiones permite establecer procedimientos, patrones y criterios precisos. Lembcke¹⁰ ha hecho gran hincapié en la necesidad de seleccionar algunas dimensiones específicas de atención dentro de determinadas clases de diagnóstico, en lugar de tratar de hacer evaluaciones generales de dimensiones inespecíficas que, en su opinión, carecen de exactitud. Emplea las siguientes dimensiones: confirmación del diagnóstico clínico, justificación del tratamiento (incluso cirugía) e integridad del procedimiento quirúrgico. Dentro de cada dimensión y en cada clase de diagnóstico, se usan a menudo una o más actividades previamente definidas para clasificar los resultados correspondientes a esa dimensión en su conjunto. Son ejemplos de ello la compatibilidad del diagnóstico de pancreatitis con las concentraciones de amilasa sérica o la del diagnóstico de cirrosis del hígado con los resultados de una biopsia, la práctica de pruebas de sensibilidad antes de la antibioterapia en casos de bronquitis aguda y el control del nivel de glucemia en la sangre en casos de diabetes.

Además de la medida en que se realiza la selección previa de las dimensiones, la evaluación de la calidad difiere con respecto al número de dimensiones usadas y la minuciosidad con que se exploran los resultados en cada dimensión. Por ejemplo, Peterson y colaboradores¹⁸ y Rosenfeld²² emplean un gran número de dimensiones. Por otra parte, Peterson y Barsamian^{38,39} se concentran en dos dimensiones básicas, a saber, justificación del diagnóstico y de la terapia, pero exigen prueba completa de la justificación. Huntley *et al.*⁴³ dan un ejemplo de un método más simplificado al evaluar la atención ambulatoria con dos criterios solamente: el porcentaje de exámenes que excluyen ciertos procedimientos de rutina y el porcentaje de anomalías encontradas cuya trayectoria no se siguió.

Cualquier juicio emitido respecto de la calidad es incompleto cuando se usan sólo algunas dimensiones y cuando las decisiones sobre cada una se adoptan basándose en pruebas parciales. Algunas dimensiones, como los cuidados preventivos o la atención psicológica y social de la salud y la enfermedad, se excluyen a menudo de la definición de calidad y de los patrones y criterios que le permiten funcionar. Son ejemplos de ello la exclusión intencional de la atención psiquiátrica del estudio de Peterson¹⁸ y la de la relación entre el paciente y el médico y la actitud de los médicos en los estudios sobre la calidad de la atención en el Plan de Seguro Médico de la Zona Metropolitana de Nueva York²⁷. Rosenfeld²² insistió especialmente en incluir los resultados de determinadas medidas exploratorias entre los criterios de atención superior; pero la atención se consideró buena sin esas medidas. A falta de instrucciones específicas a los jueces, en el estudio de Morehead *et al.*²⁶ se incluyen historias de casos que, en su opinión, recibieron óptima atención, en los que cualquier falla en materia de atención preventiva podría haber tenido graves consecuencias para el paciente.

Otra característica de medición es el nivel en que se fija el patrón. Los patrones pueden ser tan estrictos que nadie cumpla con ellos o tan tolerantes que todos reciban una "buena" clasificación. Por ejemplo, en el estudio de la práctica general notificado por Clute¹⁹, el examen de la tensión arterial, la medición de la temperatura corporal, la otoscopia y la aplicación de inmunizaciones no sirvieron para clasificar a todos los médicos porque todos ellos tuvieron un buen desempeño.

Escalas de medición

La capacidad de hacer una distinción entre diferentes grados de desempeño depende de la escala de medición usada. En muchos estudios de calidad se emplea un número mínimo de divisiones para clasificar la atención, considerada como un conjunto, en categorías como "excelente", "buena", "regular" o "mala". La posición relativa de una persona en cada conjunto se puede especificar más computando el porcentaje de casos en cada categoría de la escala. En otros estudios se asigna cierta calificación a los resultados de determinados elementos de la atención y se acumulan los valores respectivos para obtener un índice numérico que, por lo general, va de 0 a 100. Esas prácticas dan lugar a interrogantes relativos a las escalas de medición y al funcionamiento legítimo de las mismas. A continuación se describen algunas de ellas.

Quienes defienden la primera práctica señalan que no es posible lograr un mayor grado de precisión con los métodos

actuales. Algunos han llegado a reducir las clases a sólo dos: óptima y subóptima. Clute¹⁹ emplea tres y reconoce que la del medio es dudosa o indeterminada. Además, la atención médica tiene un aspecto de totalidad o nulidad que no es posible captar en la calificación numérica normal. La atención puede ser buena en muchos aspectos y desastrosamente inadecuada en su conjunto por un error fundamental de uno de sus elementos. Por supuesto, este no es un problema común, si se demuestra que los resultados obtenidos con diversos elementos de atención guardan una estrecha correlación.

Quienes han empleado el sistema de calificación numérica han señalado que se pierden mucha información al emitir un juicio general³⁸ y que la calificación numérica, acumulada a partir de determinada calificación secundaria, da una idea no sólo del conjunto, sino de la evaluación de cada parte. Rosenfeld²² ha resuelto este problema valiéndose de un sistema de asignación de una calificación cualitativa particular a los elementos de la atención y de una global basándose en reglas arbitrarias de combinación que dan cabida al atributo de totalidad o nulidad de la calidad de la atención médica. Como ya se señaló, hubo mucha concordancia entre la clasificación intuitiva y estructurada en el estudio de Rosenfeld²² y entre la cualitativa y cuantitativa en el de Peterson *et al.*¹⁸.

Un grave problema, todavía sin resolver, en la formulación de un sistema de calificación numérica, es la forma de ponderación de los distintos elementos en el proceso de cálculo del total. En la actualidad, ese es un asunto arbitrario. Por ejemplo, Peterson *et al.*¹⁸ formulan la siguiente escala: historia clínica 30, examen físico 34, uso de ayudas de laboratorio 26, terapia 9, medicina preventiva 6, fichas clínicas 2, total 107. Daily y Morehead²⁴ asignan diferentes grados de importancia, de la manera siguiente: fichas 30, exámenes de diagnóstico 40, tratamiento y seguimiento 30, total 100. Peterson *et al.* dicen que "se asigna la mayor importancia al proceso de emisión de un diagnóstico, ya que sin éste la terapia no puede ser racional. Además, la terapia se encuentra en proceso de cambio constante, mientras que la forma de tomar la historia y practicar el examen físico se ha modificado muy poco en el correr del tiempo"¹⁸. Daily y Morehead no ofrecen ninguna justificación de sus índices de ponderación, pero tal vez se podrían presentar en su nombre argumentos igualmente persuasivos. Sigue existiendo el problema de la búsqueda de confirmación externa⁴⁴.

El problema de los índices de ponderación está relacionado con el de naturaleza más general del valor de los puntos de información o de procedimientos en el proceso de atención médica. Rimoldi *et al.*³⁴ emplearon la frecuencia con que se usaron determinados puntos de información para solucionar un problema de ensayo como medida del valor de dicho punto. Williamson hizo que los expertos clasificaran determinados procedimientos en un medio específico para la práctica de pruebas de diagnóstico, en una escala que osciló entre "muy útil" y "muy nocivo". Los resultados obtenidos en cada prueba se clasificaron luego empleando índices cuantitativos de "eficiencia", "pericia" y "competencia" general, según la frecuencia y naturaleza de los procedimientos empleados³⁵.

Un problema encontrado en la interpretación de la calificación numérica es el significado del intervalo numérico de los

puntos de la escala. La calificación numérica asignada con fines de evaluación de la calidad generalmente carece de intervalos iguales. No se debe usar como si tuviera esa propiedad.

Confiabilidad

La confiabilidad de las evaluaciones es un punto de gran importancia en los estudios de calidad, en los que tantas cosas dependen del juicio emitido, aun cuando se usan patrones detallados. En varios estudios se ha prestado cierta atención al acuerdo entre los jueces. Se tiene la impresión de que se considera aceptable. Basándose en otras 14 visitas hechas por observadores, Peterson *et al.*¹⁸ consideraron que había suficiente acuerdo para poder agrupar todas las observaciones en una de las seis clases principales de atención, después de hacer las modificaciones necesarias par tener en cuenta el sesgo de parte de cada observador. En el estudio de Daily y Morehead "se hicieron varias verificaciones del trabajo de los dos médicos internistas entrevistadores, pidiéndoles que entrevistarán a los mismos médicos. Las diferencias en la calificación dada a los médicos de familia basándose en la clasificación asignada por separado no fueron superiores a 7%"²⁴. Rosenfeld²² prestó gran atención a los análisis de fiabilidad de las pruebas y formuló índices matemáticos de "acuerdo" y "dispersión" para medirla. Esos índices muestran que existe bastante acuerdo, pero es difícil hacer una evaluación precisa puesto que no se ha comprobado que ningún otro investigador haya usado esas mismas medidas. Morehead *et al.*²⁶, en el segundo estudio de la atención médica que recibieron los familiares de los miembros del sindicato de conductores de camión, indican que hubo acuerdo inicial entre dos jueces al asignar la atención a una de dos clases en 78% de los casos. Esa cifra aumentó a 92% después de reevaluar los casos de desacuerdo de los dos jueces.

En contraste con la confiabilidad de la evaluación entre jueces, poco se ha dicho sobre la de los juicios emitidos respecto de la calidad en repetidas ocasiones por la misma persona. Para someter a ensayo la variación del juicio de un mismo observador, Peterson *et al.*¹⁸ les pidieron a dos de ellos que volvieran a visitar a cuatro de los médicos que habían visitado antes. El grado de concordancia del juicio de un mismo observador fue menor que el existente entre uno y otro, en parte porque las nuevas visitas fueron más breves y, por tanto, guardaron relación con una muestra de menor tamaño.

El principal mecanismo para lograr un mayor grado de fiabilidad es la especificación detallada de criterios, patrones y procedimientos empleados para evaluar la atención. De hecho, el esfuerzo por duplicar el proceso representó un gran impulso para el establecimiento de los sistemas de clasificación más rigurosos de Lembcke y de Peterson y Barsamian. Lamentablemente, no existen estudios comparativos de confiabilidad entre los métodos de evaluación muy específicos en sus indicaciones y los que no lo son. Los datos sin elaborar de Rosenfeld podrían permitir una comparación de la fiabilidad de los juicios "intuitivos" y la de los juicios estructurados de los mismos dos evaluadores. Se podrían analizar los datos inéditos de Morehead *et al.*²⁶ de la misma manera que los de Rosenfeld²² para dar información útil sobre la relación existente entre el grado de fiabilidad y el método de evaluación. Los datos parciales que se han publicado sugieren que la fiabilidad lograda por Morehead *et al.* después

del examen, al emplear el método menos indicativo, es similar a la lograda por Rosenfeld, que empleó una técnica mucho más indicativa.

Morehead y colaboradores preguntaron si la fiabilidad proveniente de la especificación detallada de patrones y criterios no se logra a costa de una menor validez, lo que reviste importancia. "A menudo, esos criterios obligan a introducir en un rígido marco acciones o factores similares que quizá no sean apropiados en una situación dada, debido a las infinitas variaciones de la reacción del cuerpo humano a la enfermedad...". El grupo de estudio rechaza la hipótesis de que esos criterios sean necesarios para evaluar la calidad de la atención médica. Opinan unánimemente que la flexibilidad es tan importante para los evaluadores al juzgar cada caso como para el médico competente al confrontar el problema clínico de una paciente determinado²⁶.

Las razones del desacuerdo entre los jueces aclaran los problemas de evaluación y las posibilidades de lograr mayor confiabilidad. Rosenfeld descubrió que "casi la mitad de las diferencias eran atribuibles a situaciones no cubiertas debidamente por los patrones o en las que estos eran ambiguos. En otro ámbito surgieron diferencias sobre preguntas reales, porque un consultor pasó por alto un punto de información importante en el registro. Por tanto, al parecer, se podría lograr un grado de concordancia mucho mayor con patrones revisados y mejores métodos de orientación de los consultores²². Menos del 25% de los casos de desacuerdo se refieren a diferencias de opinión en lo relativo a los requisitos para el manejo de los pacientes. Esta es una función de ambigüedad del sistema de atención médica y fija un límite máximo de duplicación. Morehead *et al.* informan que en cerca de la mitad de los casos de desacuerdo inicial "hubo concordancia en cuanto al aspecto más serio de la atención del paciente, pero un evaluador afirmó más tarde que no había tenido en cuenta los corolarios de esa atención²⁶. Otras razones del desacuerdo fueron la dificultad para adherirse a diversas escalas de clasificación o el no haber tenido en cuenta todos los hechos. Del reducido número de casos de desacuerdo sin resolver (8% de todos los pacientes internados y 36% de los casos de desacuerdo inicial) más de la mitad se debieron a diferencias de opinión bien intencionadas respecto del manejo clínico del problema. El resto surgió de diferencias de interpretación de registros inadecuados o de los problemas técnicos relativos al punto en que se debe evaluar la atención insatisfactoria en una serie de hospitalizaciones²⁷.

Un último aspecto de la confiabilidad es el fallo ocasional en el desempeño de un evaluador, como se ha demostrado en forma tan impresionante en el estudio de Reosenfeld²². Se planteará aquí la pregunta, sin intención de contestarla, de qué es lo que hace el investigador cuando un segmento bien definido de sus resultados aparece totalmente aberrante.

Sesgo

Cuando varios observadores o jueces describen y evalúan el proceso de atención médica, uno de ellos puede emplear constantemente patrones más rígidos que otro o interpretar patrones previamente determinados en forma más estricta. Peterson *et al.*¹⁸ descubrieron que uno de sus observadores daba generalmente una clasificación más alta que el otro al evaluar

la práctica del examen físico, pero no a la prestación de otras clases de atención. Rosenfeld²² mostró que uno de los dos evaluadores asignaba regularmente una clasificación más baja a los mismos casos estudiados por ambos. El examen de cada caso de desacuerdo en el estudio hecho por Morehead y colaboradores²⁶ revela que, en la clase correspondiente a atención médica, el mismo evaluador dio a la atención una clasificación más baja en 11 de 12 casos de desacuerdo. En lo que respecta a casos quirúrgicos, un evaluador asignó una clasificación más baja a la atención que el otro en los ocho casos de desacuerdo. Al examinar las razones de desacuerdo en los casos médicos se tiene la impresión de que uno de los jueces tenía especial interés en la cardiología y exigía más claridad y seguridad en el tratamiento de los casos de cardiopatía.

Estos resultados indican claramente que hay que aceptar el sesgo como la regla más que la excepción y que los estudios de calidad se deben formular teniendo eso en cuenta. Por ejemplo, en el estudio de Rosenfeld²² cualquiera de los dos evaluadores cuyos servicios se emplearon en cada campo de ejercicio profesional habrían calificado a varios hospitales en el mismo orden, aunque uno fue siempre más generoso que el otro. Por otra parte, el estudio del ejercicio general hecho por Clute¹⁹, en Canadá, se ha criticado por comparar la calidad de la atención en dos regiones geográficas, aunque diferentes observadores examinaron la atención en las dos zonas en cuestión⁴⁵. El autor conocía ese problema y encontró métodos de comparación del desempeño de los observadores en las dos zonas geográficas, pero la deficiencia básica permanece igual.

El orden o la regularidad previamente determinados en el proceso de estudio puede guardar relación con el sesgo. Por tanto, es posible que haya que introducir procedimientos cuidadosamente planeados al formular investigaciones con fines de aleatorización. El estudio realizado por Peterson *et al.*¹⁸ parece ser uno de los pocos en los que se ha prestado atención a ese factor. Otra importante fuente de sesgo está en los conocimientos que tenga el evaluador respecto de la identidad del médico que prestó la atención o del hospital en que se dispensó. Se ha preguntado si es posible eliminar las características de identificación de los cuadros objeto de análisis³, pero poco se sabe sobre la factibilidad de ese procedimiento y sus efectos en la clasificación asignada. Puede producirse todavía otra clase de sesgo como resultado de los restringidos patrones y criterios de práctica que pueden surgir en ciertas instituciones o "facultades" de ejercicio de la medicina y a su alrededor. Mientras eso sea verdad o se sospeche de que sea, habrá que tomar las debidas precauciones para la consecución y asignación de jueces.

Validez

La eficacia de la atención, como se dijo, para lograr o producir salud y satisfacción, según la definición de una sociedad o subcultura particular para cada uno de sus miembros, es el elemento fundamental de comprobación de la validez de la calidad de la misma. La validez de los demás fenómenos como indicadores de la calidad depende, en último análisis, de la relación existente entre esos fenómenos y el logro de salud y satisfacción. Sin embargo, la conformidad de la práctica con los patrones aceptados tiene una clase de validez condicional o

parcial que puede ser más pertinente para fines de evaluación en determinados casos.

La confirmación de la validez de los detalles de la práctica de la medicina por sus efectos para la salud es de particular interés para las ciencias clínicas. En las publicaciones sobre medicina clínica se buscan datos para determinar si la penicilina fomenta la recuperación en ciertas clases de neumonía, los anticoagulantes en la trombosis coronaria o los corticosteroides en la carditis reumática; lo que indican ciertos análisis sobre la función hepática; y si la mastectomía simple o radical es el procedimiento que más prolonga la vida en determinadas clases de cáncer de mama. De los conocimientos generales sobre esas relaciones surgen los patrones de la práctica, cuya validez se ha comprobado plenamente en mayor o menor grado y mediante los cuales se juzga de ordinario el proceso de atención médica.

Los extremos de naturaleza intermedia o relativos a procedimiento representan a menudo mayores conjuntos de atención. Su relación con el resultado ha atraído la atención del investigador clínico y del estudiante de la organización de atención médica. Algunos ejemplos de esto último son los estudios de las relaciones existentes entre la atención prenatal y la salud maternoinfantil^{46,47} y entre los exámenes exploratorios múltiples y el estado de salud subsiguiente⁴⁸. Un interesante ejemplo del estudio de la relación existente entre un extremo relativo a procedimiento y otro es el intento de demostrar que existe una relación favorable entre la práctica de los exámenes rectales y vaginales por el médico y la confirmación patológica de apendicitis en apendectomía primarias, según se informa en el Estudio de las Actividades Profesionales⁴⁹.

En muchos estudios reseñados^{18,19,23,26,28} se trata de analizar la relación existente entre las propiedades estructurales y la evaluación del proceso de atención. Por ejemplo, en varios de estos estudios se ha demostrado que existe una relación entre la formación e idoneidad de los médicos y la calidad de la atención que prestan. Sin embargo, la relación es compleja y recibe la influencia de la clase de formación, su duración y el tipo de hospital en que se recibió. Los dos estudios de la práctica general^{18,19} han mostrado otras relaciones favorables entre la calidad y la disponibilidad de mejores consultorios para la práctica, la dotación de equipo de laboratorio y la institución de un sistema de registro de citas. No se demostró que hubiera ninguna relación entre la calidad y el pertenecer a asociaciones profesionales, el ingreso del médico o la dotación de equipo de radiografía del consultorio. Los dos estudios no concuerdan completamente respecto de la naturaleza de la relación existente entre la calidad de la práctica y la formación del médico en un hospital docente o en otro, el número de horas trabajadas o la naturaleza de la afiliación hospitalaria del médico. La acreditación de los hospitales, presuntamente una marca de calidad conferida principalmente por la observación de una amplia gama de patrones de organización, en sí y por su propia naturaleza, no parece guardar relación con la calidad de la atención, al menos en la ciudad de Nueva York²⁶.

Aunque, sin duda alguna, la estructura y el proceso están relacionados, los pocos ejemplos citados indican claramente la complejidad y ambigüedad de esas relaciones. Este es el resultado, en parte, de los muchos factores que intervienen y, en

parte, de interacciones mal entendidas de los mismos. Por ejemplo, basándose en varios resultados^{26,38}, se podría proponer razonablemente que los factores relativos a los hospitales y a los médicos influyen en la calidad de la atención prestada en el hospital, pero que las diferencias entre los médicos se eliminan en los mejores y los peores hospitales y se expresan, en distintos grados, en los de calidad intermedia.

Un método preferido por los estudiantes de organización de la atención médica consiste en examinar las relaciones existentes entre la estructura y el resultado, sin referirse a los complejos procesos que los vinculan. Ya se han citado algunos ejemplos de esos estudios⁶⁻⁹. Otros incluyen estudios de los efectos que tiene la reorganización de una clínica de atención ambulatoria para el estado de salud⁵⁰, los efectos de la atención hospitalaria intensiva en la recuperación⁵¹, los efectos de la atención en el hogar en la supervivencia⁵² y el efecto de un programa de rehabilitación en el estado físico de los pacientes atendidos en sanatorios particulares^{53,54}. La falta de una relación con el resultado en los dos últimos estudios sugiere que, en algunos casos, la opinión actual sobre la forma de organizar la atención no está bien definida.

En esta breve reseña se indican las clases de pruebas relativas a la validez de los diferentes métodos de evaluar la calidad de la atención. Obviamente, no se entiende a cabalidad la relación entre el proceso y el resultado ni entre estos dos últimos y la estructura. Con respecto a ello, los requisitos para confirmar la validez se pueden expresar mejor con el concepto, ya citado, de una cadena de acontecimientos en la que cada uno marca el final del precedente y representa una condición necesaria para el siguiente. Eso indica que la relación de los medios con el fin entre pares adyacentes exige que se confirme la validez en cualquier cadena de acontecimientos hipotéticos o reales⁵⁵. Por supuesto, este es un proceso laborioso. Como se ha demostrado, los vínculos entre uno y otro son los que más comúnmente se pasan por alto. Como resultado, las inferencias causales se atenúan en proporción de la distancia que separa a los dos acontecimientos en la cadena.

Lamentablemente, existe muy poca información sobre la evaluación real de la calidad en la que se emplea más de un método de evaluación simultáneamente. Los estudios de Makover se han concentrado en particular en la relación existente entre la evaluación multifactorial de la estructura y la del proceso en los mismos grupos médicos. "Se descubrió que los grupos médicos que lograron clasificaciones altas en cuanto a calidad por el método usado en el presente estudio fueron los que, en general, se ciñeron más estrictamente a los patrones médicos mínimos fijados por el Plan de Seguro Médico. Sin embargo, las excepciones fueron suficientemente notorias, en número y grado, para poner en tela de juicio la fiabilidad⁵⁶ de uno u otro método de clasificación cuando se aplicó a cualquier grupo médico. Al parecer, es obvio que se necesita hacer otras comparaciones de estos dos métodos de clasificación"²³.

Índices de atención médica

Puesto que una evaluación multidimensional de la atención médica es una empresa costosa y laboriosa, se siguen buscando datos específicos fácilmente mensurables que propor-

cionen información sobre la calidad de la atención médica. Los datos usados pueden referirse a varios aspectos de estructura, proceso o resultado. El principal requisito es que se puedan medir con facilidad, a veces, en forma regular, y que tengan una validez razonable. Ente los estudios de calidad en los que se emplea este método están el de actividades profesionales³⁶ y los de Ciocco⁵⁷ y Furstenberg *et al.*³⁷.

Esos índices tienen la ventaja de ser convenientes; pero las inferencias derivadas de los mismos pueden ser de dudosa validez. Myers ha señalado las muchas limitaciones de los índices tradicionales de la calidad de la atención hospitalaria, incluso tasas de mortalidad total y postoperatoria, complicaciones, infección postoperatoria, cesárea, consultas y extirpación de tejido normal en una operación⁵⁸. La precisión e integridad de la información básica puede estar sujeta a dudas. Más importante aún es la posibilidad de que surjan serios interrogantes sobre lo que significa cada índice, ya que son muchos los factores que intervienen para que se produzca el fenómeno que mide. Por otra parte, Eislee ha señalado que por lo menos ciertos índices pueden ser útiles, si se emplean con cuidado³⁶.

La búsqueda de formas fáciles de medir un fenómeno tan complejo como la atención médica puede ser un sueño imposible. El uso de índices sencillos en lugar de medidas complejas se puede justificar demostrando una elevada correlación entre los mismos¹. No obstante, por no haberse podido demostrar que existen vínculos causales, ésta puede ser una base inestable como punto de partida. Por otra parte, cada índice puede ser una medida de una dimensión o un elemento de atención. Por tanto, los índices múltiples seleccionados acertadamente pueden constituir el equivalente de perforaciones de reconocimiento en un estudio geológico, que dan suficiente información sobre las partes para poder reconstruir el todo. Por supuesto, la validez de las inferencias sobre el todo dependerá de la medida de la continuidad interna en el ejercicio particular o institucional de la medicina.

Algunos problemas para evaluar la atención ambulatoria

Ya se han mencionado algunas de las dificultades especiales para evaluar la calidad de la atención ambulatoria. Incluyen la escasez de información registrada y los conocimientos previos del médico encargado en lo relativo a los antecedentes médicos y sociales del paciente. El primero de esos problemas ha llevado a usar los servicios de observadores adiestrados y el segundo, a observar casos en que los conocimientos previos no constituyen un factor de importancia para el tratamiento en curso. También se ha puesto en tela de juicio el grado de importancia que para el ejercicio general tienen los patrones y estrategias de atención establecidos por médicos centrado en los hospitales y orientados hacia el ámbito académico.

Otro problema es la dificultad de definir el segmento de la atención que puede ser el objeto de evaluación en materia de atención ambulatoria⁵⁹. Cuando se trata de atención hospitalaria, una solo ingreso es comúnmente la unidad apropiada⁵⁹. Cuando se trata del ejercicio en un consultorio particular o una clínica, una secuencia de atención puede cubrir un número indeterminado de consultas, de modo que la identificación de la unidad apropiada representa un interrogante. Por lo común, la

respuesta consiste en escoger un período en forma arbitraria para definir el episodio de atención pertinente. Ciocco y colaboradores⁵⁷ lo definieron como la primera consulta más 14 días de seguimiento. Huntley *et al.*⁴³ usan un período de cuatro semanas después del examen inicial.

Conclusiones y propuestas

En esta reseña se ha tratado de dar una idea de los diversos enfoques y métodos empleados para evaluar la calidad de la atención médica y de señalar ciertos asuntos y problemas destacados por los mismos, que exigen consideración.

Cabe afirmar que los métodos empleados han sido de dudoso valor y carecen a menudo de rigor y precisión. Pero, ¿cuán precisas tienen que ser las estimaciones de calidad? Al menos los mejores métodos han sido adecuados para los fines de política administrativa y social que les han dado origen. La búsqueda de perfección no debería cegarnos respecto del hecho de que las actuales técnicas de evaluación de la calidad, por rudimentarias que sean, han revelado una escala de calidad que va de sobresaliente a deplorable. Hoy en día se dispone de medios para emitir juicios de amplio alcance con bastante seguridad. Este grado de seguridad se apoya en resultados a los que ya se ha hecho alusión, que sugieren que existe un grado aceptable de homogeneidad en la práctica individual y de poder duplicar los juicios cualitativos basados en un método de evaluación de estructuración mínima. Eso no quiere decir que no quede todavía mucho por hacer para lograr la mayor precisión que exigen ciertos otros fines.

Se podría comenzar un catálogo de los puntos que se necesita perfeccionar, considerando la naturaleza de la información que constituye la base de cada juicio emitido respecto de la calidad. Se debe saber más sobre el efecto que tiene el observador en la práctica analizada y sobre el proceso de observación propiamente dicho y su fiabilidad y validez. Es preciso hacer comparaciones entre la observación directa y la información registrada con y sin el complemento de una entrevista con el médico encargado. No basta registrar los casos en que hay acuerdo o desacuerdo. Se necesita un estudio más detallado de la naturaleza y las razones de la discrepancia en varios ámbitos. Asimismo, debe someterse a prueba el uso de resúmenes de fichas frente al de las fichas mismas.

El proceso de evaluación propiamente dicho exige un estudio más detallado. Se dedica un gran esfuerzo a la formulación de criterios y patrones que presuntamente dan estabilidad y uniformidad a cualquier juicio emitido sobre la calidad y, sin embargo, ese presunto efecto no se ha demostrado empíricamente. No se sabe hasta dónde debe ir la normalización explícita para lograr adelantos considerables en materia de fiabilidad. También se debe considerar si con una mayor normalización se pierde tanto la capacidad de dar cuanta de los elementos imprevistos en la situación clínica que se logra fiabilidad a costa de validez. La evaluación del mismo conjunto de fichas con patrones y criterios cada vez más estructurados debe dar información valiosa sobre estos puntos. La afirmación de que los evaluadores menos idóneos que usan criterios minuciosos pueden emitir un juicio fidedigno y válido también se puede someter a prueba de esa forma.

Anteriormente se ha indicado lo poco que se sabe respecto de la fiabilidad y del sesgo cuando se comparan dos o más jueces y de la fiabilidad de los juicios emitidos en repetidas ocasiones sobre los mismos puntos de atención por el mismo evaluador. Asimismo, se sabe muy poco sobre los efectos que en la fiabilidad y validez tienen ciertas características de los jueces, incluso su experiencia, sus campos de particular interés y los factores de su personalidad. Se puede aprender mucho sobre estos y otros asuntos afines haciendo explícito el proceso de emisión de juicios y sometiéndolo a un cuidadoso estudio. Eso debe revelar las dimensiones y los valores empleados por los diversos jueces y mostrar cómo se resuelven las diferencias cuando dos o más jueces discuten sus puntos de vista. Hoy en día existe alguna duda respecto de la validez de los actos de conciliación realizados en grupo en los que puede predominar un punto de vista, no necesariamente por ser más válido¹. El efecto de enmascarar la identidad del hospital o del médico que presta atención se puede estudiar de la misma manera. Aquí se propone no sólo que se demuestren las diferencias o similitudes en cuanto a juicios en general sino que, al hacer explícita la forma de pensar de los jueces, se trata de determinar cómo surgen las diferencias y similitudes y cómo se resuelven las primeras.

Además de los defectos del método, en la mayoría de los estudios se ha adoptado una definición de calidad demasiado limitada. En general, se ocupan del tratamiento técnico de la enfermedad y prestan poca atención a la prevención, rehabilitación, coordinación y continuidad de la atención o al manejo de la relación entre el paciente y el médico. Presuntamente, la razón de ello se centra en que los requisitos técnicos del tratamiento se reconocen más ampliamente y están más normalizados. Por tanto, se necesita una exploración conceptual y empírica más completa de la definición de calidad.

El significado de "exploración conceptual" se puede explicar considerando la dimensión de eficiencia que se pasa por alto a menudo en los estudios de calidad. Se podrían distinguir dos tipos de eficiencia: lógica y económica. La eficiencia lógica tiene que ver con el uso de información para llegar a una decisión. En este caso, el problema podría ser si la información obtenida por el médico es pertinente o no para el asunto clínico que será objeto de transacción. En caso afirmativo, se podría considerar el grado de repetición o duplicación de los datos obtenidos y la medida en que excede de los requisitos de la adopción de decisiones en una situación dada. Si la parsimonia representa un valor en la atención médica, la identificación de redundancia se convierte en un elemento de la evaluación de la atención.

La eficiencia económica se refiere a las relaciones entre insumos y productos y lleva a preguntarse si se puede obtener un determinado producto a menor costo. Por supuesto, recibe la influencia de la eficiencia lógica, ya que la acumulación de información innecesaria o no utilizada es un procedimiento costoso que no reporta ningún beneficio. Típicamente, se extiende más allá del individuo y guarda relación con el producto social del esfuerzo de atención médica. Entraña consideración de la posibilidad de que la "mejor" atención médica para el individuo tal vez no lo sea para la comunidad. Peterson y colaboradores citan un ejemplo que resume este asunto. "Dos médicos habían

delegado la supervisión de las consultas prenatales ordinarias en las enfermeras del consultorio y uno de ellos atendía a la paciente sólo si tenía alguna queja particular"¹⁸. Por un lado, esta puede haber sido atención subóptima para cada madre embarazada. Por otro, puede haber sido una brillante estrategia para proporcionar la pericia conjunta de un equipo de atención médica al mayor número posible de mujeres. Cordero, en un trabajo que incita a la reflexión, ha documentado la tesis de que, cuando los recursos son limitados, la atención médica óptima a la comunidad puede exigir un nivel subóptimo para cada uno de sus miembros⁶⁰.

Además de la exploración conceptual del significado de calidad, en función de las dimensiones de atención y de los valores que se les asignan, se necesitan estudios empíricos de lo que son las dimensiones y los valores predominantes en grupos de población pertinentes⁵. Por ejemplo, poco se sabe respecto de la forma en que los médicos definen la calidad y no se conoce la relación ente el ejercicio del médico y su propia definición de calidad. Este es un campo de investigación importante tanto para la educación médica como para la calidad. Los estudios empíricos del proceso de atención médica también deben contribuir mucho a la identificación de las dimensiones y los valores que se deben incorporar en la definición de calidad.

Una reseña de los estudios de calidad muestra una cierta repetitividad desalentadora en los conceptos, enfoques y métodos básicos. Cualquier otro adelanto sustantivo, fuera del perfeccionamiento de la metodología, es factible que provenga de un programa de investigación del proceso de atención médica propiamente dicho más bien que de ataques directos al problema de la calidad. Se cree que eso es así porque, antes de poder emitir un juicio respecto de la calidad, se necesita entender la interacción del paciente con el médico y la actuación de este en el proceso de prestación de atención. Una vez que se entiendan los elementos del proceso y su interrelación, se puede juzgar su valor en función de su aporte al logro de metas intermedias y finales. Suponga, por ejemplo, que el conjunto formado por el autoritarismo y la permisividad representa una dimensión de la relación del paciente con el médico. Un estudio empírico puede mostrar que los médicos se diferencian, de hecho, por ese atributo. Se podría preguntar entonces si el autoritarismo o la tolerancia debe ser el criterio de calidad. La respuesta podría derivarse de los valores generales de la sociedad que puede apoyar a uno u otro como el atributo más deseable en las interacciones sociales. Esta es una forma de emitir un juicio respecto de la calidad y es perfectamente válida, siempre y cuando su razón justificativa y sus bases sean explícitas. Sin embargo, el estudio del proceso de atención médica en sí puede ofrecer un enfoque distinto y más pragmático. Suponga, de momento, que el cumplimiento de las recomendaciones del médico es una meta y un valor en el sistema de atención médica. El valor del autoritarismo o la permisividad se puede determinar, por su aporte al cumplimiento. La validez de ese cumplimiento en sí se debe confirmar con el criterio de los resultados de la atención, que es de orden superior. Tal vez el verdadero estado de cosas sea más complejo que el ejemplo hipotético dado. Se puede dar el caso de que el criterio de calidad sea la congruencia con las esperanzas de los pacientes o una adaptación más compleja a determinadas situaciones

clínicas y sociales, en lugar del autoritarismo o la permisividad como modalidad predominante. Además, tal vez ciertas metas del proceso de atención médica no sean compatibles con otras y uno no pueda referirse a la calidad en términos generales sino dentro de dimensiones y fines específicos. Por tanto, la evaluación de la calidad no resultará en un juicio sumario sino en un perfil complejo, según lo ha sugerido Sheps¹.

Por supuesto, una gran parte de la investigación sobre el proceso de atención médica se centrará en la manera en que los médicos acopian información de importancia clínica y adoptan decisiones en cuanto a diagnóstico y terapia. Este no es el lugar apropiado para presentar un marco conceptual de investigación de esta parte del proceso de atención médica. Sin embargo, se pueden mencionar determinados estudios y dar algunas indicaciones para realizar otras investigaciones.

Las investigaciones sobre el acopio de información incluyen estudios sobre la percepción e interpretación de los signos físicos^{61,62}. Por ejemplo, Evans y Bybee han demostrado que, al interpretar los sonidos cardíacos, se cometieron errores de percepción (del ritmo y la frecuencia), junto con otros de interpretación sobre lo percibido. El diagnóstico equivocado, juzgado en comparación con un criterio, fue el resultado de estos dos errores⁶². Eso señala la necesidad de incluir en las estimaciones de calidad información sobre la fiabilidad y validez de los datos producto de las funciones de los sentidos, sobre los que se basa parcialmente el tratamiento.

El trabajo de Peterson y Barsamian^{38,39} representa el método más cercano a una rigurosa evaluación de las decisiones adoptadas en materia de diagnóstico y terapia. Como tal, es quizá el adelanto reciente de mayor importancia en cuanto a métodos de evaluación de la calidad. No obstante, ese método se basa en reseñas de fichas y se destina casi exclusivamente a justificar el diagnóstico y la terapia. Como resultado, en la evaluación se excluyen muchas dimensiones importantes de la atención. Algunas de estas son asuntos relativos a eficiencia y estilos y estrategias para la resolución de problemas.

Los estilos y estrategias para la resolución de problemas se pueden estudiar por medio de la observación real de la práctica, como lo hicieron con tanta eficacia Peterson y colaboradores en su estudio de la práctica general¹⁸. Una gran parte de lo que no se ha observado puede hacerse explícito al pedirle al médico que diga en voz alta lo que está haciendo y por qué. Este método de *reflexión en voz alta* se ha empleado en estudios de resolución de problemas aunque, en sí, puede modificar la conducta⁶³. Otro método consiste en crear situaciones experimentales, como las establecidas por Rimoldi *et al.*³⁴ y por Williamson³⁵ para observar el proceso de adopción de decisiones. Aunque esas situaciones tienen ciertas limitaciones que emanan de su artificialidad⁶⁴, su mayor simplicidad y el mayor control que ofrecen pueden ser de gran utilidad.

A primera vista, el estudiante de atención médica podría esperar ayuda de los conocimientos teóricos y prácticos provenientes del campo general de la investigación en lo que respecta a solución de problemas. Lamentablemente, no existe una base teórica bien establecida que se pueda explotar con facilidad en los estudios de atención médica. Algunos de los estudios empíricos en materia de solución de problemas podrían sugerir métodos e ideas aplicables a situaciones de aten-

ción médica⁶³⁻⁶⁷. Ciertos estudios de resolución de problemas relativos a equipo electrónico, en particular, muestran una intrigante similitud con el proceso de diagnóstico y tratamiento médico. En estos y en otros similares se han identificado características de comportamiento que se podrían emplear para clasificar diversos estilos de tratamiento clínico. Incluyen el volumen de información recolectada, la intensidad de la búsqueda de información, el valor de los puntos de información buscados y modificados según su lugar en una secuencia e interacción con otros, varias clases de redundancia, el surgimiento de estereotipos, los patrones de investigación en relación con la parte defectuosa, las tendencias a obrar antes de conseguir suficiente información o a buscarla más allá del punto en que se puede garantizar razonablemente una solución, la "distancia con respecto a los errores", el éxito en el logro de una solución y así sucesivamente.

La teoría de adopción de decisiones también puede ofrecer instrumentos conceptuales de investigación en el proceso de atención médica. Ledley y Lusted^{68,69}, entre otros, han tratado de aplicar modelos basados en probabilidades condicionales al proceso de diagnóstico y terapia. Peterson y Barsamian^{38,39} decidieron no emplear probabilidades en sus sistemas lógicos por no tener a su disposición los datos necesarios (la probabilidad independiente de enfermedades y síntomas y la de determinados síntomas de ciertas enfermedades). Sin embargo, Edwards y colaboradores⁷⁰ señalan que todavía se puede someter a prueba la eficiencia en materia de adopción de decisiones empleando los datos estadísticos que uno preferiría tener, en sustitución de la probabilidad subjetiva (la de la persona que adopta las decisiones o la de expertos seleccionados).

Una pregunta básica que surge con frecuencia a lo largo del presente artículo es la medida en la cual la actuación en materia de atención médica es un fenómeno homogéneo o heterogéneo. Por ejemplo, se observó que esto era de importancia para el muestreo, el uso de índices en lugar de medidas multidimensionales y la construcción de escalas con las que se pretende juzgar la actuación en su conjunto. Cuando se hace esta pregunta con respecto a médicos individuales, el objeto del estudio es la integración de varias clases de conocimientos y habilidades relativos a la personalidad y la conducta del médico. Cuando se formula esa pregunta respecto de las instituciones y los sistemas sociales, los factores son completamente distintos. En este caso uno se preocupa por los mecanismos formales e informales para organizar, moldear y dirigir el esfuerzo humano, en general, y la práctica de la medicina, en particular. Se espera que la investigación en todos estos campos contribuya a una mayor complejidad en lo que respecta a la medición de la calidad.

Algunas de las convenciones aceptadas en la presente reseña constituyen, por su propia naturaleza, obstáculos para un estudio más importante de la calidad. En el mundo real, los servicios de los médicos no se prestan aparte de los de otros profesionales de salud ni de los de una gran variedad de funcionarios de apoyo. La separación de la atención hospitalaria y ambulatoria también es bastante artificial. Las unidades de atención que son los objetos de estudio propiamente dichos incluyen los aportes de muchas personas durante una secuencia que puede incluir atención en varios ámbitos. La forma en que se definen e identifican esas secuencias tiene repercusiones para

el muestreo, los métodos de acopio de información y los patrones y criterios de evaluación.

Un último comentario se relaciona con la actitud con que se abordan los estudios sobre calidad. Ya se han mencionado los imperativos sociales que llevan a evaluar la calidad. A menudo guardan relación con ellos el celo y los valores del reformador social. En los estudios de calidad se necesita mayor neutralidad y objetividad. A menudo hay que preguntarse "¿qué sucede aquí?" en lugar de "¿hay algo que anda mal?" y "¿cómo se puede mejorar?". Eso no quiere decir que el investigador renuncie a sus propios valores u objetivos sociales, sino que se reconoce y mantiene la distinción entre los valores y los elementos de estructura, proceso o resultado y que ambos están sujetos a un estudio igualmente crítico. En parte, para lograr esa clase de orientación, hay que dejar de preocuparse por evaluar la calidad par concentrarse en entender el proceso de atención médica propiamente dicho.

Nota de agradecimiento

Entre los autores, aquí incluidos, que leyeron la versión preliminar e hicieron correcciones o comentarios están Georgopoulos, Makover, Morehead, Peterson, Riedel, Rosenstock, Rosenfeld, Sheps y Weirnerman. El autor tiene una profunda deuda de gratitud con la Dra. Mildred A. Morehead y con los profesores Basil S. Georgopoulos, Herbert E. Klarman y Charles A. Metzner por haber tomado tiempo para hacer extensos comentarios. El Sr. Sam Shapiro y el Dr. Jonas N. Muller, críticos oficiales, fueron de gran ayuda en la definición precisa de algunos problemas de evaluación de la calidad. Puesto que el autor no pudo emplear en su totalidad los excelentes consejos recibidos, asume plena responsabilidad por cualquier falla del presente documento.

Bibliografía y notas*

1. Sheps MC. Approaches to the quality of hospital care. *Public Health Reports* 1955;70:877-86.

Este documento representa una materialización insólitamente fructífera de las ideas relativas a la evaluación de la calidad. Contiene discusiones breves pero sumamente completas de los fines de evaluación, los problemas de definición, los criterios y patrones, los distintos sistemas de medición, la fiabilidad de los juicios cualitativos y los índices de calidad. La bibliografía es excelente.

2. Peterson L. Evaluation of the quality of medical care. *New England Journal of Medicine* 1963;269:1238-45.
3. Lerner M, Riedel DC. The teamster study and the quality of medical care. *Inquiry* 1964;1:69-80.

El valor principal de este documento está en que formula preguntas sobre métodos de evaluación, incluido el muestreo de poblaciones y las diversas clases de diagnóstico, el uso de fichas y la necesidad de suplementación con entrevistas, el valor de los patrones detallados, la necesidad de entender el proceso de verificación, la definición de términos y conceptos (por ejem-

plo, de "hospitalización innecesaria") y los problemas de definición del episodio de atención pertinente.

4. Lee RI, Jones LW. *The Fundamentals of Good Medical Care*. Chicago. University of Chicago Press, 1933.
5. Klein MW, et al. Problems of measuring patient care in the outpatient department. *Journal of Health and Human Behavior* 1961;2:138-44.
6. Kohl SG. *Perinatal Mortality in New York City: Responsible Factors*. Cambridge, Harvard University Press, 1955.

Este estudio, patrocinado por la Academia de Medicina de Nueva York, fue un examen de las fichas relativas a una muestra representativa de defunciones perinatales en la ciudad de Nueva York, hecho por un comité de expertos. Se reconocieron las defunciones prevenibles y se identificaron los "factores de responsabilidad", incluidos los errores de juicio y técnica cometidos por el médico. La incidencia de esos errores tuvo que ver también con el tipo de servicio hospitalario, servicio profesional y hospital e indicó la relación existente entre la estructura y el resultado, modificada por las características de la población servida.

7. Shapiro S, et al. Further observations on prematurity and perinatal mortality in a general population and in the population of a prepaid group practice medical care plan. *American Journal of Public Health* 1960;50:1304-17.
8. Lipworth L, Lee JAH, Morris JN. Case fatality in teaching and nonteaching hospitals, 1956-1959. *Medical Care* 1963;1:71-6.
9. Rice CE, et al. Measuring social restoration performance of public psychiatric hospitals. *Public Health Reports* 1961;76:437-46.
10. Lembcke PA. Medical auditing by scientific methods. *Journal of the American Medical Association* 1956;162:646-55. (Los apéndices A y B fueron suministrados por el autor).

Éste es quizá el documento en que mejor se describen los conceptos básicos y métodos del sistema de un alto grado de estructuración formulado por Lembcke para revisar las fichas de los hospitales. También se incluye un ejemplo del extraordinario efecto que puede tener una "revisión externa" de esta clase en la práctica quirúrgica en un hospital.

11. Kelman HR, Willner A. Problems in measurement and evaluation of rehabilitation. *Archives of Physical Medicine and Rehabilitation* 1962;43:172-81.
12. McDermott W, et al. Introducing modern medicine in a navajo community. *Science* 1960;131:197-205 y 280-7.
13. Simon HA. *Administrative Behavior*. Nueva York: The Macmillan Company, 1961;62-6.
14. Hutchinson GB. Evaluation of preventive services. *Journal of Chronic Diseases* 1960;11:497-508.
15. James G. *Evaluation of Public Health. Report of the Second National Conference on Evaluation in Public Health*. Ann Arbor: Universidad de Michigan, Facultad de Salud Pública, 1960;7-17.
16. Weirnerman ER. Appraisal of medical care programs. *American Journal of Public Health* 1950;40:1129-34.
17. Goldmann F, Graham EA. *The Quality of Medical Care Provided at the Labor Health Institute, St. Louis, Missouri*. St. Louis: The Labor Health Institute, 1954.

Éste es un buen ejemplo de un método de evaluación basado en características estructurales. En este caso, estas incluyen la disposición y el equipo de las instalaciones físicas, la competencia y

*Las referencias bibliográficas se reproducen en su formato original con las notas del autor

estabilidad del personal médico, las disposiciones tomadas con fines de continuidad del servicio centrado alrededor del médico de familia, la programación y duración de las consultas clínicas, el asunto tratado en el examen inicial, la atención prestada a la medicina preventiva y la idoneidad de las fichas médicas.

18. Peterson OL, et al. An analytical study of North Carolina general practice: 1953-1954 (2ª parte). *The Journal of Medical Education* 1956;31:1-165.

Este estudio, que ya es clásico, se distingue por la extraordinaria atención prestada a los métodos y por la minuciosa exploración de la relación existente entre la clasificación de la calidad y las características de los médicos, incluso su formación y sus métodos de práctica. Los resultados de éste y de otros estudios en los que se ha empleado el mismo método llevan a formular preguntas básicas sobre la práctica general que se considera tradicional en este y en otros países.

19. Clute KF. *The General Practitioner: A Study of Medical Education and Practice in Ontario and Nova Scotia*. Toronto: University of Toronto Press, 1963, capítulos 1, 2, 16, 17 y 18.

Puesto que este estudio se basa en el método formulado por Peterson y colaboradores, ofrece una excelente oportunidad de examinar la naturaleza general de la relación entre las características de los médicos y la clasificación de la calidad. Además, el lector de este volumen elegantemente escrito puede disfrutar de una descripción detallada de la práctica general en los dos campos de estudio.

20. Kroeger HG, et al. The office practice of internists, I. The feasibility of evaluating quality of care. *The Journal of the American Medical Association* 1965;193:371-6.

Ésta es la primera de una serie de documentos basados en el estudio de la práctica de los afiliados a la Sociedad de Medicina Interna de Nueva York. En este documento se presentan resultados relativos a la integridad de las fichas de los consultorios, su idoneidad para juzgar la calidad y el grado de concordancia entre los resúmenes de fichas preparados por los médicos y por otro personal bien adiestrado. No se presentan los juicios emitidos respecto de la calidad de la atención prestada. Otros documentos de esta serie publicados en el *Journal of the American Medical Association* tienen que ver con el número de pacientes (23 de agosto), sus características (13 de septiembre), las actividades profesionales distintas de la atención de pacientes particulares (11 de octubre) y los antecedentes y la forma de ejercicio (1 de noviembre).

21. Kilpatrick GS. Observer error in medicine. *Journal of Medical Education* 1963;38:38-43. Véase una bibliografía útil sobre los errores de los observadores en Witts LJ, ed. *Medical Surveys and Clinical Trials*. Londres: Oxford University Press, 1959;39-44.

22. Rosenfeld LS. Quality of medical care in hospitals, *American Journal of Public Health* 1957;47:356-65.

Este estudio comparativo de la calidad de la atención en cuatro hospitales, cuidadosamente formulado, se centra en los problemas de los métodos de evaluación de la calidad. Ofrece importante información sobre el uso de patrones normativos y empíricos, la fiabilidad y el sesgo de los juicios basándose en un análisis por medio de cuadros, la correlación de los defectos del registro con los de la práctica y la homogeneidad de la clasificación de la calidad en cada clase de diagnóstico y entre una y otra.

23. Makover HB. The quality of medical care: Methodological survey of the medical groups associated with the health insurance plan

of New York. *American Journal of Public Health* 1951;41:824-32.

Éste es quizá el primer informe publicado respecto de un programa de estudios de la calidad de la atención en grupos médicos que han suscrito algún contrato con el Plan de Seguro Médico de la Zona Metropolitana de Nueva York, que tiene una estructura administrativa pero está orientado hacia la investigación. Lamentablemente, gran parte de este trabajo se mantiene inédito. Una característica particular de este documento está en que describe y presenta los resultados de una evaluación simultánea de la estructura (políticas, organización, administración, finanzas y actividades profesionales) y el proceso (evaluación de una muestra de fichas clínicas).

24. Daily EF, Morehead MA. A method of evaluating and improving the quality of medical care. *American Journal of Public Health* 1956;46:848-54.

25. Fitzpatrick TB, Riedel DC, Payne BC. Character and effectiveness of hospital use. En: Mc Nerney WJ et al., *Hospital and Medical Economics*. Chicago, Hospital Research and Education Trust, American Hospital Association, 1962;495-509.

26. Morehead MA, et al. *A Study of the Quality of Hospital Care Secured by a Sample of Teamster Family Members in New York City*. Nueva York: Universidad de Columbia, Facultad de Salud Pública y Medicina Administrativa, 1964.

Este estudio y otro de la misma índole²⁸ realizan una función social y administrativa de suma importancia al documentar los casos en que a menudo resulta ser inadecuada la atención recibida por los miembros de un sindicato por medio de fuentes tradicionales. Estos estudios ayudan mucho a comprender las relaciones existentes entre las características de los hospitales y los médicos y la calidad de la atención que prestan. Se tienen en cuenta diversas clasificaciones de los médicos por campo de especialización y privilegios de internado de pacientes, así como la clasificación de los hospitales por dueño, afiliación a una facultad de medicina, autorización para ofrecer formación en residencia y acreditación. Los efectos interactivos de algunas de estas variables también son objeto de estudio. Además, el segundo de los dos estudios²⁶ se centra en asuntos de metodología, incluso el muestreo representativo de los casos de hospitalización frente al muestreo crítico y la fiabilidad de la evaluación de las fichas por diferentes jueces.

27. Morehead MA. Comunicación personal.

28. Ehrlich J, Morehead MA, Trusell RE. *The Quantity, Quality and Costs of Medical and Hospital Care Secured by a Sample of Teamster Families in the New York Area*. Nueva York: Universidad de Columbia, Facultad de Salud Pública y Medicina Administrativa, 1962.

29. Maloney MC, Trusell RE, Elinson J. Physicians choose medical care: A sociometric approach to quality appraisal. *American Journal of Public Health* 1960;50:1678-86.

Este estudio representa un ingenioso método de evaluación por medio de los "juicios emitidos por iguales" en lo que se cree que es una situación particularmente reveladora: la clase de atención que se escoge para el médico o sus propios familiares. Algunas de las características de los médicos y cirujanos seleccionados incluyeron antiguas relaciones personales y profesionales, especialización reconocida y afiliación a una facultad de medicina. Una valiosa información incidental indica que si bien 9 de cada 10 médicos declararon que toda persona debe tener su propio médico, 4 de cada 10 dijeron que tenían a alguien que consideraban

su médico particular y sólo 2 de cada 10 habían consultado a su médico particular en el transcurso del último año.

30. Georgopoulos BS, Mann FC. *The Community General Hospital*. Nueva York: The Macmillan Company, 1962.

El estudio de la calidad notificado en varios capítulos de este libro se basa en la tesis de que si uno desea averiguar cuál es la calidad de la atención prestada, todo lo que hay que hacer es preguntárselo a las personas que la prestan, de manera directa o indirecta. Aunque los médicos pueden considerar que esta es una idea bastante ingenua, la estabilidad y coherencia interna de los resultados notificados en el presente estudio indican que este método merece consideración más cuidadosa. En un segundo estudio de una muestra nacional de hospitales generales se tratará de confirmar la validez de las opiniones de los entrevistados al compararlas con índices seleccionados de actividades profesionales en cada hospital. Los resultados se esperarán con gran interés.

31. El Sr. Arnold D. Kaluzny, uno de los estudiantes del autor, le ayudó a acuñar esta palabra.
32. Georgopoulos BS, Tannenbaum AS. A study of organizational effectiveness. *American Sociological Review* 1957;22:534-40.
33. Evans LR, Bybee JR. Evaluation of student skills in physical diagnosis. *Journal of Medical Education* 1965;40:199-204.
34. Rimoldi HJA, Haley JV, Fogliatto H. *The Test of Diagnostic Skills*. Loyola Psychometric Laboratory, Publicación N° 25, Chicago: Loyola University Press, 1962.

Este estudio es de interés porque se centra en una situación en que se realizan experimentos controlados para estudiar el desempeño de los estudiantes de medicina y de los médicos. Más interesante aún es el intento por abordar la cuestión del valor o la utilidad de las medidas en materia de diagnóstico de una forma sistemática y rigurosa. Si bien este estudio en particular no parece contribuir mucho a la comprensión de la calidad de la atención, al parecer vale la pena seguir este método general.

35. Williamson JW. Assessing clinical judgment. *Journal of Medical Education* 1965;40:180-7.

Éste es otro ejemplo de la evaluación de la actuación clínica tomando como punto de partida una situación basada en un experimento artificial. El aspecto del trabajo más digno de mencionarse es la asignación de cierto valores ("útil" o "nocivo") a un conjunto de medidas de diagnóstico y terapia y la formulación de medidas de "eficiencia", "pericia" y "competencia", basándose en las actividades seleccionadas por el sujeto para abordar el caso experimental. Con este método se detectaron diferencias de desempeño entre un médico y otro. Un resultado inesperado fue la falta de diferencias sistemáticas por edad, formación o clase de ejercicio en los grupos estudiados hasta ahora.

36. Eislee CW, Slee VN, Hoffmann RG. Can the practice of internal medicine be evaluated? *Annals of Internal Medicine* 1956;44:144-61.

Los autores discuten el uso de índices de los cuales se podrían hacer inferencias sobre la calidad del tratamiento quirúrgico y médico. Los índices descritos incluyen informes de patología de los tejidos en casos de apendectomía, pacientes diabéticos cuya concentración sanguínea de azúcar no se ha determinado y a quienes no se han tomado radiografías y casos de neumonía en los que no se ordenaron radiografías de pulmones. Un resultado sorprendente notificado en este documento, y en otros basa-

dos en el mismo método, se centra en la tremenda variación de la frecuencia de esos índices de "actividad profesional" por médico y hospital.

37. Furstenberg FF, et al. Prescribing as an index to quality of medical care: A study of the Baltimore City medical care program. *American Journal of Public Health* 1953;43:1299-309.
38. Peterson OL, Barsamian EM. *An Application of Logic to a Study of Quality of Surgical Care*. Documento presentado en el Quinto Simposio Médico de la IBM, Endicott, Nueva York, 7-11 de octubre de 1963.

En este documento y otro de la misma índole³⁹ se presenta una descripción bastante completa del método basado en el "árbol lógico" para la evaluación de la calidad. Se dan ejemplos de los sistemas lógicos empleados para el estudio del síndrome de Stein-Leventhal y de fibromioma uterino. No se ofrecen datos sobre los resultados empíricos en cuya consecución se ha empleado este método.

39. Diagnostic Performance. En: Jacquez JA, ed. *The Diagnostic Process*. Ann Arbor: The University of Michigan Press, 1964; 347-62.
40. Lembcke PA, Johnson OG. *A Medical Audit Report*. Los Angeles: Universidad de California, Facultad de Salud Pública, 1963 (documento mimeografiado).

Este documento representa una extensión del método de revisión de la atención médica de Lembcke a las diferentes clases de diagnóstico médico y a un gran número de intervenciones quirúrgicas. Aunque este volumen es un compendio de datos sin elaborar en su mayoría, un cuidadoso estudio puede ofrecer información sobre el método empleado por el autor y sus limitaciones.

41. Lembcke PA. A scientific method for medical auditing. *Hospitals* 1959;33:65-71(16 de junio);65-72(1 de julio de 1959).
42. Está por determinarse la dimensión del conjunto de variables en las que se incorporan esos patrones.
43. Huntley RR, et al. The quality of medical care: Techniques and investigation in the outpatient clinic. *Journal of Chronic Diseases* 1961;14:630-42.

Este estudio ofrece un ejemplo de la aplicación de un procedimiento ordinario de análisis por medio de cuadros para verificar la calidad de la atención prestada en el departamento ambulatorio de un hospital docente. Con bastante frecuencia se dejaron de realizar los procedimientos ordinarios y se determinó que no se habían seguido los casos de anomalías. Al parecer, un procedimiento modificado de análisis por medio de cuadros ocasionó una importante reducción del porcentaje de anomalías excluidas del seguimiento.

44. Peterson, et al., loc. cit., trataron de confirmar los índices de ponderación por medio del procedimiento de análisis de los factores. Los que son de importancia matemática se citan en su nota de las págs. 14-15.
45. Mainland D. *Calibration of the Human Instrument*. Notes from a Laboratory of medical Statistics, 1964;81 (documento mimeografiado).
46. *Comité Mixto del Real Colegio de Obstetricia y Ginecología y el Comité de Investigaciones sobre Población*. Maternity in Great Britain. Londres: Oxford University Press, 1948.

47. Yankauer A, Goss KG, Romero SM. An evaluation of prenatal care and its relationship to social class and social disorganization. *American Journal of Public Health* 1953;43:1001-10.
48. Wylie CM. Participation in a multiple screening clinic with five-year follow-up. *Public Health Reports* 1961;76:596-602.
49. Comisión de Actividades Profesionales y Hospitalarias. *Medical Audit Study Report 5: Primary Appendectomies*. Ann Arbor: Comisión de Actividades Profesionales y Hospitalarias, octubre de 1957.
50. Simon AJ. Social structure of clinics and patient improvement. *Administrative Science Quarterly* 1959;4:197-206.
51. Lockward HJ, Lundberg GAF, Odoroff ME. Effect of intensive care on mortality rate of patients with myocardial infarcts. *Public Health Reports* 1963;78:655-61.
52. Bakst JN, Marra EF. Experiences with home care for cardiac patients. *American Journal of Public Health* 1955;45:444-50.
53. Muller JN, Tobis JS, Kelman HR. The rehabilitation potential of nursing home residents. *American Journal of Public Health* 1963;53:243-7.
54. Estos estudios también incluyen datos sobre las relaciones existentes entre las características estructurales y los extremos relativos a procedimiento. Son ejemplos de ello el efecto de la estructura clínica en el número de consultas ambulatorias⁵⁰ y el efecto de un programa de atención en el hogar en la hospitalización⁵².
55. Getting VA, et al. *Research in Evaluation in Public Health Practices*. Documento presentado en la 92ª Reunión Anual de la Asociación Americana de Salud Pública, Nueva York, 5 de octubre de 1964.
56. Suponiendo que el criterio empleado sea la evaluación directa del proceso, el asunto se convierte en una de las repercusiones que tienen las medidas de fiabilidad en la validez.
57. Ciocco A, Hunt H, Altman I. Statistics on clinical services to new patients in medical groups. *Public Health Reports* 1950;65:99-115.
- Ésta es una aplicación preliminar a la práctica de grupo del análisis de "actividades profesionales" relacionado hoy en día generalmente con la evaluación de la atención hospitalaria. Los índices usados incluyeron el registro del diagnóstico y tratamiento, la práctica de exámenes rectales y vaginales, la práctica de ciertos análisis de laboratorio y el uso de sedantes, estimulantes y otros medicamentos que son causa de uso indebido. Como sucede en los hospitales, los grupos variaron mucho con respecto a esos indicadores.
58. Myers RS. Hospital statistics don't tell the truth. *Modern Hospital* 1954;83:53-4.
59. Aún en el caso de la atención hospitalaria, la unidad apropiada puede incluir la atención prestada antes y después de la hospitalización, así como varias hospitalizaciones³.
60. Cordero AL. The determination of medical care needs in relation to a concept of minimal adequate care: An evaluation of the curative outpatient services in a rural health center. *Medical Care* 1964;2:95-103.
61. Butterworth JS, Reppert EH. Auscultatory acumen in the general medical population. *Journal of the American Medical Association* 1960;174:32-4.
62. Evans LR, Bybee JR. Evaluation of student skills in physical diagnosis. *Journal of Medical Education* 1965;40:199-204.
63. Fattu NC. Experimental studies of problem solving. *Journal of Medical Education* 1964;39:212-25.
64. John ER. Contributions to the study of the problem solving process. *Psychological Monographs* 1957;71.
65. Duncan CP. Recent research in human problem solving. *Psychological Bulletin* 1959;56:397-429.
66. Fattu NA, Mech E, Kapos E. Some statistical relationships between selected response dimensions and problem-solving proficiency. *Psychological Monographs* 1954;68.
67. Stolurow LM, et al. The efficient course of action in "trouble shooting" as a joint function of probability and cost. *Educational and Psychological Measurement* 1955;15:462-77.
68. Ledley RS, Lusted LB. Reasoning foundations of medical diagnosis. *Science* 1959;130:9-21.
69. Lusted LB, Stahl WR. *Conceptual Models of Diagnosis*. En: Jacquez JA, ed. *The Diagnostic Process*. Ann Arbor: The University of Michigan Press, 1964;157-74.
70. Edwards W, Lindman H, Phillips LD. *Emerging Technologies for Making Decisions*. En: Newcomb TM, ed. *New Directions in Psychology, II*. Nueva York: Holt, Rinehart & Winston, Inc., 1965;261-325.

Criterios y métodos de evaluación de la calidad asistencial

Donabedian A. The quality of medical care. *Science* 1978;200:856-64

La calidad de la atención médica se enfrenta a un panorama de indefiniciones conceptuales, de carencias metodológicas e incluso de resistencias abiertas. Los intentos por evaluar dicha calidad se han enfrentado a estructuras de poder gremial sustentadas en recursos y mitos como la ignorancia del público en relación a los procesos mórbidos y su incapacidad para definir sus necesidades.

La mayor parte de las evaluaciones de la atención médica se ha limitado a un recuento de las acciones emprendidas, sin plantear su efecto sobre la salud ni la medida en que satisfacen las necesidades de los pacientes. En el fondo de estas omisiones subyace el asunto de la calidad.

El artículo empieza con la definición y los criterios de la calidad. Donabedian define la calidad como "una adecuación precisa de la atención a las necesidades particulares de cada caso", con una evidente implicación para los servicios colectivos. A lo largo del artículo aparece una reiterada preocupación por una real democratización de la calidad de la atención, vista en términos de su aportación a la duración y a la calidad de la vida.

La segunda parte del artículo emprende una revisión completa de los métodos de evaluación de la calidad de la atención de la cual destacan los hallazgos de las fallas en el sistema de atención médica. Adicionalmente, Donabedian se aboca a una tarea original de ordenación y clasificación de dichos métodos.

La última sección explica el contexto del monitoreo de la calidad en un país industrializado: los Estados Unidos. El análisis de Donabedian sobre los tipos de fuerzas que llegan a producir una intervención gubernamental en la monitorización de una

atención en su mayor parte privada reviste, sin duda, un interés general. Será conveniente observar la evolución de ese experimento, representado por las "Organizaciones para la Revisión de las Normas Profesionales".

Enfrentados aún a vastas carencias cuantitativas, muchos países han soslayado la cuestión de la calidad de los servicios de salud. Sin embargo, ya es tiempo de cuestionarse no solamente por la cantidad sino por el tipo y calidad de servicios que se necesitan. La ampliación de los servicios se ha realizado principalmente sobre la base de un modelo de atención que restringe sus fundamentos teóricos al las dimensiones biológicas e individuales, olvidando el panorama epidemiológico, las necesidades y las demandas explícitas de la mayoría de la población.

Por ello, el análisis del problema de la calidad debe empezar con una definición del modelo de atención que se ajuste a las necesidades de salud, para proseguir con la determinación precisa de normas de calidad. El acceso franco a la atención es un componente de la calidad de la vida, por lo que será necesario garantizar la calidad de la misma.

El artículo de Avedis Donabedian ofrece los elementos para emprender una profunda reflexión a este respecto. En él podrá encontrarse, esa doble dimensión enriquecedora del estado del arte y de la sistematización estimulante en tanto integra una lectura profunda capaz de descubrir lo que permanece implícito.

Julio Frenk
Secretario de Salud. México

La Calidad de la Atención Médica*

Avedis Donabedian

Hemos concedido a las profesiones de la salud acceso a los sitios más secretos y sensibles de nosotros mismos, y les hemos confiado asuntos que conciernen a nuestro bienestar, felicidad y supervivencia. A cambio, hemos esperado que las profesiones se gobiernen a sí mismas de un modo tan estricto que no debamos temer explotación o incompetencia.

El objetivo de la evaluación de la calidad es determinar con qué grado de éxito se ha conseguido esto; y el propósito de la monitorización de la calidad es el ejercer una vigilancia constante para que cualquier alejamiento de las normas pueda ser detectado tempranamente y corregido. Pero primero, debemos especificar qué es lo que se está evaluando y monitorizando.

Definición de la calidad de la atención

La definición de la calidad de la atención implica la especificación de tres factores: 1. el fenómeno que constituye el objeto de interés, 2. los atributos del fenómeno sobre el cual se va a emitir un juicio, 3. los criterios y las normas que coloquen a cada atributo en una escala que vaya de lo mejor posible a lo peor. Como existe una considerable falta de claridad respecto a cada uno de estos tres elementos, el problema de la evaluación de la calidad está lleno de desacuerdos y de confusión. Con relación al fenómeno objeto de interés, la mayor división ocurre entre dos escuelas de pensamiento. La primera se limita al desempeño de un profesional o técnico de la salud -o bien un grupo de ellos relacionados entre sí de una manera funcional- en la atención de las personas que se han convertido en sus pacientes. La segunda presenta una visión más amplia e insiste en que el objeto de interés es un programa extenso o, incluso, un sistema de atención médica. Según este punto de vista, al juzgarse la calidad de la atención debe evaluarse no sólo lo que sucede a los que la reciben, sino también cuantas personas se encuentran privadas de ella. La evaluación también debe determinar si la distribución de atención médica, entre quienes la necesitan, permite que tanto los individuos como las poblaciones reciban el tipo de atención del cual se pueden esperar beneficios óptimos. En otras palabras, el acceso y otros aspectos de la distribución de los recursos, así como los efectos externos, han llegado a convertirse en características fundamentales para evaluar la calidad. Sin embargo, hay que señalar que estas características no están totalmente au-

tes de la perspectiva más restringida, ya que los factores que influyen en el acceso a la atención lo hacen también en la posibilidad de continuar siendo un paciente dispuesto a cooperar. Además, el problema de la distribución de recursos está relacionado con la forma en que los profesionales y técnicos distribuyen su tiempo entre los múltiples pacientes que demandan atención.

En el presente artículo me ceñiré a una visión un tanto restringida, basada principalmente en el desempeño de los médicos en su manejo de pacientes individuales. Respecto a los atributos sobre los cuales habrá de emitirse un juicio, se distinguen dos aspectos. Uno es la aplicación del conocimiento médico y de la tecnología de forma que se obtenga un máximo de beneficios con un mínimo de riesgos, teniendo en cuenta las preferencias de cada paciente. El otro aspecto es el manejo de la relación personal con el paciente, que debe ser conforme a los principios éticos, a las convenciones sociales y a las expectativas y necesidades legítimas del paciente.

No debe pensarse que dicha división en dos aspectos corresponde a la distinción entre la "ciencia" y el "arte" de la medicina. El balance de riesgos y beneficios de los procedimientos médicos en cada caso particular es, en sí, tanto una ciencia como un arte, dependiendo de cómo de explícitas y bien comprendidas sean las operaciones mentales que las determinan. De forma similar, el manejo de la relación interpersonal es en gran medida un "arte, debido básicamente a que no ha sido estudiado tan sistemáticamente como requeriría. De cualquier modo, tanto la aplicación de la ciencia y la tecnología de la medicina, como el manejo de la relación interpersonal, son elementos fundamentales de la calidad de la atención médica, sobre todo desde el punto de vista del paciente. Por desgracia, en los estudios sobre la calidad de la atención se ha hecho hincapié en la primera y se ha tendido a ignorar la segunda, de modo que gran parte de lo que describiré a continuación será parcial.

Un atributo -el del coste- es tan importante para la política social que merece ser mencionado por separado. Tradicionalmente, al tratar las especificaciones abstractas de la calidad de la atención médica, se ha excluido su coste, como si esto no fuera un obstáculo y cualquier paciente tuviera acceso a todo cuanto la medicina puede ofrecer. Si bien esta posición tiene un sólido fundamento ético, ignora ciertas realidades importantes. Todo el mundo coincide en que no se debe hacer nada que no

*Donabedian A. The quality of medical care. *Science* 1978;200:856-64

Reproducción autorizada por los propietarios del copyright para la traducción española en homenaje a la figura del profesor Donabedian. Versión española supervisada por Fundación Avedis Donabedian y Revista de Calidad Asistencial

aporte beneficios al paciente o que implique más riesgos que beneficios. Tales maniobras son costosas y denotan una falta de criterio del médico. Pero supongamos que únicamente se prescriben aquellos procedimientos de los que se espera un beneficio neto, según el nivel actual de conocimientos.

Cuando escasea el dinero, es lógico suponer que se usarán aquellos procedimientos en los que el beneficio neto resulte muy alto en relación al coste. Pero conforme aumenta la cantidad de dinero disponible, la atención se vuelve más elaborada y se van agregando procedimientos que proporcionan beneficios netos exiguos en relación al coste. Se llega a un punto en el cual, si se explicara la situación con claridad a un paciente que pagara por su propia atención, éste se negaría a someterse a más maniobras, después de haber decidido que el exiguo beneficio adicional no corresponde al coste adicional.

Según la definición de calidad que se ha adoptado en este artículo, el médico tendría la obligación de informar al paciente y a su familia del balance de riesgos y beneficios y del coste monetario del beneficio neto esperado, a fin de que fuera posible tomar una decisión conjunta sobre lo que debe hacerse y sobre el momento de detenerse; en función de estas consideraciones cabe esperar que la decisión varíe de un caso a otro. En términos más generales, este planteamiento lleva a la conclusión de que a cada incremento de la "calidad" corresponde un coste monetario, y de que se espera que los pacientes que consultan al médico decidan en qué momento la "calidad" adicional no tiene el valor del coste adicional y por lo tanto no se desea, lo cual equivale a decir que ya no se le puede llamar "calidad". Todo esto supone que cada paciente paga por sí mismo los costes y recibe el beneficio completo de la atención médica, suposición que en nuestro complejo mundo, no es válida.

Los seguros de enfermedad y los programas gubernamentales reparten el coste de la atención entre mucha gente. Los beneficios de la atención pueden extenderse a otras personas además de aquellas que la reciben directamente, y la sociedad puede hacer mayor hincapié en atender la salud de ciertos sectores -los niños, por ejemplo- que la de otros. Esto significa que la decisión social sobre los casos que la calidad adicional no corresponde al coste adicional que puede diferir de la decisión del individuo.

De ahí que el médico tenga que enfrentarse a los intereses y deseos del paciente individual y las obligaciones que la sociedad le impone. El deseo, comprensible, de evadir este dilema moral puede explicar, en parte, la resistencia que han mostrado los médicos ante los intentos de imponer normas públicas para la calidad de la atención. Por desgracia, el médico parece estar condenado a que su trabajo se encuentre siempre ceñido a este dilema. En tiempos pasados, debía abstenerse de hacer lo que consideraba mejor para el paciente, cuando éste no podía sufragar los gastos; ahora quizá deba abstenerse de realizar algo que pudiera ser útil al paciente porque la sociedad lo limita. Pero antes, si sufría con esta situación, lo hacía en privado, mientras que hoy en día teme ser atacado públicamente.

Criterios y estándares de calidad

Con el fin de evaluar, la definición de la calidad debe hacerse en términos precisos y funcionales, a manera de criterios y normas específicos. Al hacerlo, nos enfrentamos a un proble-

ma fundamental. Si la calidad es una adecuación precisa de la atención a las necesidades particulares de cada caso, ¿es posible formular especificaciones pormenorizadas de lo que constituye la calidad aplicable a grupos de casos? La mayoría de los médicos responderían negativamente. Insistirían en que una evolución definitiva de la calidad de la atención debe basarse en el conocimiento de todas las circunstancias de cada caso, de manera que un asesor reconocido por su habilidad superior pueda utilizar su propio criterio como norma de comparación, al reconstruir mentalmente el criterio que hubiera seguido para la atención del caso en esas circunstancias.

Dichas evaluaciones, que recurren a los llamados criterios "implícitos", resultan, desde luego, extremadamente largas y costosas. Son también poco fiables, a menos de que las lleven a cabo médicos sumamente competentes, motivados y capaces de realizar evaluaciones. Además, la capacidad de cualquier asesor siempre puede ser cuestionada. Por ello, quienes pretenden mantener la atención médica bajo supervisión constante han recurrido al planteamiento de "criterios explícitos", que se supone representan una práctica médica a un nivel por lo menos aceptable¹. En un extremo, estos criterios y normas representan lo que los expertos más sobresalientes, basados en los mejores datos científicos, consideran como la mejor práctica; y en el otro, pueden derivarse de la práctica promedio de los médicos de una comunidad. Obviamente, se esperaría que los límites y la presunta validez de estos dos planteamientos resultaran muy diferentes y, en la práctica, se podría hacer un intento de aceptar una posición intermedia.

Independientemente del criterio que se emplee, el aspecto de la validez resulta especialmente problemático, ya que no todo en la práctica médica se acepta de forma universal, ni se encuentra apoyado sólidamente por evidencia "científica". Esto significa que hay grandes dudas y controversias, cuando menos sobre algunos de los criterios y normas en casi cualquier planteamiento. Ello explica también por qué los médicos se resisten a ser juzgados según criterios y normas ajenos a los suyos propios. Los criterios explícitos formulados de antemano dan lugar a un problema más: es difícil que dichos criterios tomen en cuenta las diferencias específicas que existen entre los distintos casos. Esto se resuelve haciendo una subclasificación de los casos más o menos homogénea y dividiendo los criterios en dos tipos, que se pueden denominar "categóricos" y "contingentes".

Los criterios categóricos son listas de maniobras que simplemente deben o no deben realizarse en todos los casos que correspondan a una clase determinada². Los criterios contingentes son listas de maniobras que pueden o deben realizarse en algunos casos, pero no en otros, dependiendo de la naturaleza y de las circunstancias de cada caso. Un mayor refinamiento consiste en especificar la frecuencia con que se espera que cada maniobra se realice en una muestra "representativa" de los casos de una determinada clase³.

La mayoría de los estudiosos del tema estaría de acuerdo en que los criterios explícitos, formulados de este modo, son recursos útiles para identificar los casos sospechosos de no cumplir los estándares y que el grado de concordancia constituye una medida aproximada de calidad. No obstante, la mayoría de los médicos no insistirá en que en ningún caso dado se

puede basar un juicio definitivo en la concordancia de los criterios que, se supone, son aplicables al "caso promedio". Aún hace falta someter todo caso de atención dudosa a un juicio por parte de los médicos expertos, a quienes se les proporcionen todos los datos relevantes y de quienes se espere que no sólo utilicen los criterios explícitos sino también el conjunto, mucho más amplio, de criterios implícitos que rigen la atención de los individuos en toda su complejidad.

De lo anterior se deduce que la mayoría de los sistemas de monitorización de la calidad de la atención se realiza en dos etapas: en la primera se identifican los casos que no cumplen los criterios explícitos; en la segunda se someten dichos casos a un examen detallado por parte de otros médicos, práctica que es conocida como "revisión por colegas" ("peer review"). Se puede recurrir a asesores externos, en lugar de los colegas, cuando haya oposición al juicio inicial, cuando una agencia externa tenga la responsabilidad inicial o de supervisión, o bien cuando se este llevando a cabo una investigación. Esta combinación de un estudio inicial seguido de un examen detallado en el seno de la organización que proporciona la atención, o fuera de ella, cumple con los objetivos de la monitorización, siempre que se tenga la voluntad y la capacidad para emplearla en forma adecuada. Sin embargo, no llena por completo los requisitos más rigurosos de un inicio válido y fiable sobre calidad de la atención. Para ello es necesario especificar al detalle las estrategias adecuadas de atención, juzgadas según los beneficios que proporcionan y los riesgos y costes que implican. Más adelante trataré este punto.

Aproximaciones a la evaluación

De lo anterior puede deducirse que la evaluación de la calidad consiste en un juicio sobre el proceso de la atención, proporcionado por los técnicos y profesionales, en forma individual o colectiva. Cuando no se cuenta con información directa sobre dicho proceso o cuando ésta resulta incompleta, se pueden deducir algunas cosas sobre el "proceso", estudiando cosas sobre el "proceso", estudiando la "estructura" o el "resultado"⁴. Se entiende aquí por "estructura" los medios materiales y sociales empleados para proporcionar la atención. Se incluyen en ellos el número, la combinación y las aptitudes del personal, así como su forma de organización y control; el espacio, el equipo y otros medios físicos, etcétera. Al evaluar la estructura se juzga si la atención se proporciona en condiciones que permiten u obstaculizan la prestación adecuada de los servicios. Las conclusiones del análisis de la estructura pueden ser seriamente cuestionadas debido a que aún se comprende poco sobre la relación entre el proceso y la estructura.

Hay bases más sólidas para emplear el análisis del "resultado" cuando se quiere determinar la calidad de la atención prestada con anterioridad.

Los resultados de la atención son, básicamente, cambios en el estado de la salud que se pueden atribuir a dicha atención. Una visión más amplia incluye los cambios en los conocimientos relacionados con la salud y en las actitudes y el comportamiento del paciente⁵. El estado de salud en sí puede ser considerado, desde una perspectiva limitada, como una función física y fisiológica; o, desde una perspectiva más amplia, que

incluye también las funciones psicológicas y el desempeño social⁶. De hecho, se están realizando numerosas investigaciones con el fin de idear alguna forma que combine estos elementos en una medida única, que refleje, no sólo la supervivencia sino también la calidad de aquella vida⁷. Si se logra idear este tipo de medida, se podría expresar la calidad de la atención en términos de su aportación a la duración y a la calidad de vida. En otras palabras, la calidad de la atención es proporcional al grado de la mejoría de la calidad de la vida que logre brindar, suponiendo que el coste no constituya un obstáculo.

En los últimos años, este planteamiento, que resulta lógico y atractivo, ha ganado numerosos adeptos e intensificado la controversia entre quienes dan prioridad a la evaluación del proceso y quienes se basan completamente en el resultado. En mi opinión, dicha controversia surge de una concepción errónea. La evaluación de la calidad no es una investigación clínica destinada a establecer la relación entre el proceso y el resultado. Es un juicio sobre el proceso de atención que se apoya en lo que ya se sabe de esa relación, tomando en cuenta los límites actuales de la ciencia médica. Es cierto que los elementos del proceso pueden emplearse como indicadores de la calidad únicamente cuando existe una relación válida entre estos elementos y los resultados deseados; también lo es que los resultados específicos pueden usarse como indicadores de la calidad sólo en la medida en que exista una relación válida entre ambos. De esta manera, la validez reside no en la selección de elementos del proceso o del resultado, sino en lo que se sabe acerca de su relación. Si existe una relación válida, puede usarse cualquiera de las dos (depende de cuál pueda ser medido con mayor facilidad y exactitud); pero si no existe, no puede usarse ninguno.

Estudios de la calidad de la atención

Las categorías de todo estudio de la calidad de la atención médica pueden ser tan variadas y la agrupación de las características tan confusa que ha resultado imposible elaborar una clasificación sencilla y satisfactoria. En el presente trabajo ignoré los estudios que dependen principalmente de evaluaciones de la estructura, y para los restantes utilizaré la clasificación de la Tabla 1.

Una breve revisión de algunos estudios tomados de esta clasificación puede ilustrar y originar preguntas sobre algunos métodos específicos de evaluación, al tiempo que proporciona información sobre algunos factores que influyen en el desempeño de las actividades de atención. Sin embargo, como algunos de estos estudios son viejos y la mayoría analiza situaciones muy limitadas, no es posible extraer conclusiones sobre niveles de calidad en general, más allá de afirmar que cada vez que la calidad de la atención ha sido examinada se encuentran graves y extensas deficiencias.

Por lo que respecta a los niveles de calidad prevalecientes en los Estados Unidos o en cualquier otro país, tenemos que confiar en medidas generales de longevidad, mortalidad, morbilidad, el uso y distribución de los servicios, la frecuencia de las intervenciones quirúrgicas y otras similares. Aunque estos aspectos son importantes, están influidos por tal cantidad de variables no examinadas que sería temerario valerse de ellos para hacer afirmaciones seguras.

Estudios del proceso de atención

La reputación de los médicos entre sus colegas surge, en gran parte, de las oportunidades que tienen para observarse unos a otros en el trabajo. El que la práctica se encuentre abierta a dicha observación constituye, en realidad, una garantía considerable, un argumento convincente a favor de la práctica organizada. Por ello resulta de especial interés descubrir el primer uso importante de una observación directa y formal con miras a la situación de la calidad del trabajo a un estudio sobre la práctica rural general, ese rincón de la medicina aislado y secreto⁸.

El método que se siguió consistió en que un médico competente, con el permiso de su anfitrión, observara a éste cuando atendía a los pacientes que iban a su primera consulta a causa de una primera enfermedad. De esta manera, fue posible para el observador juzgar si el examen físico era exhaustivo, si era apropiado realizar pruebas subsiguientes y si el tratamiento era el adecuado. Como resultado, se encontró que el 25% de los médicos cualificados como buenos o mejores que el común, mientras que en el 44% el desempeño se juzgó por debajo de un nivel "promedio" o aceptable. En general, los mejores médicos eran los más jóvenes, visitaban a sus pacientes previa cita y tenían acceso a los servicios de laboratorio; pero, sobre todo, la mayoría de ellos había tenido un periodo de entrenamiento en medicina interna, después de haberse graduado de la facultad de medicina. Todas éstas son características estructurales que facilitan una mayor atención, si bien no la garantizan. Otros estudios con el mismo enfoque sugieren que en otros países la práctica general puede presentar características y obstáculos similares⁹.

Un estudio sobre la interacción en personal de enfermería en pacientes tomados al azar en determinados hospitales del área de Detroit¹⁰, muestra que la observación de la práctica es un método con mayor aplicabilidad. Más interesantes que los niveles de desempeño encontrados, fueron los hallazgos que suponen diferencias relacionados con las características de los pacientes. Algunos aspectos de la atención de enfermería tendían a ser menos satisfactorios para los sujetos no pertenecientes a la raza blanca, para los pacientes que se encontraban en pabellones con muchas camas, para aquellos que tenían cáncer con mal pronóstico, para las mujeres jóvenes y los hombres ancianos. Debido a la naturaleza de dicho estudio, estos ejemplos no deben tomarse como definitivos, pero ilustran un problema de enorme importancia social: la medida en la que puede variar la calidad de la atención de acuerdo con las características sociales o económicas de los pacientes, ya sea porque las instituciones que proporcionan la atención sean distintas, o bien porque las mismas instituciones tengan una conducta discriminativa.

Por supuesto, la observación directa de la práctica resulta costosa y requiere de mucho tiempo. Puede también alterar la conducta observada; pero quienes han recurrido a ella afirman que pronto se olvida la presencia del observador y el individuo comienza a seguir su rutina habitual. El análisis de las historias clínicas es más discreto y se puede someter con mayor facilidad al juicio de varias personas. Sin embargo, presenta limitaciones en cuanto a que la historia clínica sea completa y veraz,

Tabla 1. Clasificación de las categorías de los estudios de calidad

- I. Estudios principalmente de la estructura
- II. Estudios principalmente del proceso
 - A. Observación directa de la práctica
 - B. Estudios Basados en el registro médico
 1. La presencia o ausencia de determinados elementos fundamentales de la atención
 2. Justificación de la cirugía y otros procedimientos de importancia
 3. Auditorias que utilizan criterios explícitos
 4. Auditorias que utilizan criterios implícitos
- III. Estudio principalmente del resultado
 - A. Morbilidad, incapacidad, mortalidad y longevidad en comunidades y poblaciones
 - B. Medidas más refinadas de morbilidad, incapacidad, mortalidad y longevidad
 1. Sucesos adversos susceptibles de prevención
 2. Progresión de enfermedades susceptibles de prevención
 3. Resultados de diagnóstico específico
 4. Mortalidad y morbilidad posoperatorias
 - C. Atribución de responsabilidades en situaciones diversas
 1. Con especificación previa de los resultados esperados
 2. Sin especificación previa de los resultados esperados
- IV. Estudios que combinan proceso y resultado para demostrar los efectos del sistema
 - A. "trayectorias"
 - B. "Rastreadores" ("tracers")
- V. Evaluación de las estrategias
 - A. Mapa de criterios
 - B. Prueba de las estrategias
 1. Con modelos
 2. Con pruebas clínicas

sobre todo en la práctica ambulatoria. Por ello se ha criticado este método en el sentido de que constituye una evaluación de los expedientes más que de la atención propiamente dicha. A ello se ha respondido que la historia clínica es un elemento importante de la atención y que existe una relación entre su calidad y la atención misma¹¹.

El análisis de la historia clínica varía en grado considerable en cuanto a la amplitud y los detalles. En un extremo, se busca sólo información sobre un pequeño número de elementos centrales que son importantes en sí y que además pueden ser tomados como representativos de ciertos aspectos de la atención que no se observan directamente. Dichos elementos centrales, o índices, pueden formularse de tal manera que resulten aplicables a todos los pacientes o subgrupos de pacientes integrados de acuerdo con la edad, el sexo, el diagnóstico, etcétera. Por ejemplo, en las historias clínicas ambulatorias se pueden investigar con que frecuencia se mide la presión arterial, se realizan exámenes rectales y vaginales, se explora el fondo del ojo y los oídos con los instrumentos adecuados, se vacuna a los niños, se realizan cultivos de estreptococos en niños que pade-

cen molestias en la garganta y exámenes de orina en las mujeres embarazadas, se prescriben sedantes, tranquilizantes y antibióticos, se aplican inyecciones en casos en que el medicamento podría haberse suministrado por vía oral, etcétera¹². Los expedientes hospitalarios permiten elaborar listas mucho más extensas de dichos indicadores, y existe una mayor seguridad de que aparecerá en el expediente la información requerida¹³. Un tipo de investigación al que se recurre con frecuencia consiste en detectar los reportes de los hallazgos anormales de laboratorio que según los médicos requieren atención y en determinar con qué frecuencia estos casos pasan desapercibidos, siendo ignorados o tratados en forma inadecuada. Por ejemplo, en la clínica general de un hospital universitario no se siguió cerca de la quinta parte de dichas alteraciones¹⁴, y en un hospital de zona mas de la mitad de los hallazgos anormales fueron ignorados o tratados de forma inadecuada¹⁵. Generalmente, cuando se reúnen los resultados de las investigaciones encaminadas a determinar los elementos fundamentales de la práctica, resulta asombroso lo variable de ésta y la frecuencia con que parece apartarse de las normas de una atención supuestamente buena.

Los avances en la recopilación y el procesamiento de los datos han estimulado el uso de este enfoque en torno a la evaluación y el seguimiento, y han ampliado su utilidad en gran medida. Los datos que aparecen en los expedientes de la consulta externa, los resúmenes de las gráficas de los pacientes hospitalizados y las demandas de pago referidas a compañías aseguradoras y programas gubernamentales, pueden alimentar un ordenador para ser procesados y confrontados con información previamente almacenada sobre el paciente, el médico o el hospital y sus divisiones. De esta manera, las aberraciones de la practica pueden identificarse, localizarse y ser examinadas más detalladamente, si su presencia o importancia lo justifica.

Además de arrojar una luz escrutadora, capaz de exponer y avergonzar al médico, el ordenador puede ser también amigo y aliado. Es posible diseñar un sistema de información que ponga al médico en estado de alerta cuando ocurran ciertos sucesos críticos determinados previamente, de manera que pueda intervenir si le parece adecuado. Como la causa aparente de muchos "errores" de la atención es la falta de ésta, más que la ignorancia, el control con ayuda informática podría constituir una importante garantía de la calidad de la atención¹⁶.

La justificación de las intervenciones quirúrgicas y de otras maniobras de importancia significan un avance desde el uso de indicadores mejorables de calidad hasta el empleo de evaluaciones más amplias y definitivas de calidad de la atención. En la misma justificación de la cirugía, se puede someter a la recomendación inicial a verificación por parte de uno o más médicos consultores; esta forma de actuar es algo que ya exigen varias compañías de seguros. En uno de esos programas los consultores estuvieron en desacuerdo con la recomendación inicial en el 18% de los casos, y hubo grandes diferencias de opinión en cuanto al diagnóstico, que oscilaban del 10% para cirugía de mama al 34% para operaciones ortopédicas¹⁷. Obviamente se puede cuestionar la mayor validez de la segunda opinión, y ésta sólo se puede verificar investigando lo que sucede a las personas que son operadas y a las que son descartadas. En un estudio de este tipo se encontró que el 30% de estas

últimas fueran operadas de todas maneras; la mitad porque desoyeron la recomendación de no operarse y la otra porque seguían presentando síntomas¹⁸.

En cuanto a aquellos que ya han sido operados, se pueden seguir dos pasos para una justificación más rigurosa. El primero consiste en determinar si el tejido extirpado se encuentra lo suficientemente afectado como para justificar la operación. Este procedimiento es tan sencillo y tan útil que se ha convertido en norma de todo hospital bien administrado.

Su validez depende, en parte de la habilidad y la integridad del patólogo, quien actúa como conciencia del hospital al sostener el espejo en el que se reflejan los fallos. Sin embargo, independientemente de la pericia con que se valore el tejido que se extirpó, la justificación de la cirugía no puede basarse únicamente en ese juicio. La decisión de operar depende del balance entre los riesgos que implica operar cuando no es necesario y los que implica no operar cuando sí es necesario; incluso la mejor de las decisiones puede traer como consecuencia la extirpación de un tejido normal. Por eso, para todo juicio definitivo debe darse un paso más allá de juzgar las condiciones del tejido extirpado, tomándose en cuenta algunas circunstancias adicionales del caso.

Una comparación entre las apendicetomías realizadas en hospitales comunitarios y hospitales docentes de Baltimore arroja luz sobre varios de estos problemas¹⁹. En los hospitales docentes, que supuestamente representan la mejor práctica, cerca de la tercera parte de los tejidos extirpados eran normales y no mostraban indicios claros de enfermedad; la proporción era la misma para los pacientes sostenidos por el gobierno que para los pacientes privados que pagaban directamente o a través de un seguro. En los hospitales comunitarios la proporción de apendicetomías en las que el tejido estaba en condiciones normales o casi normales era superior y variable según la forma en que el paciente pagara el hospital y el médico. La proporción fue del 40% para los pacientes del sistema público; del 42% para los que pagaban por sí mismos; del 50% para los que contaban con un seguro distinto al de la compañía Blue Cross, y del 55% para aquellos que contaban con seguro Blue Cross.

Aunque los resultados de este estudio no pueden ser considerados definitivos, cabe preguntarse si el hecho de contar con una protección casi total respecto coste de la atención médica, ya sea mediante un seguro privado o uno del sistema público contribuirá a elevar en gran medida las probabilidades de ser sometido a la cirugía "innecesaria". La evidencia sugiere una respuesta afirmativa. Los índices quirúrgicos varían mucho dentro de los EE.UU. y de un país a otro. Estas diferencias están relacionadas, por lo menos hasta cierto punto, con el número de cirujanos²⁰.

La cirugía también es más frecuente cuando los cirujanos trabajan solos y se les paga por operación que cuando trabajan en grupo y se les paga un salario fijo²¹. Si bien la conclusión de que se practica mucha "cirugía innecesaria" esta justificada, también es cierto que no se ha establecido por completo el índice apropiado de intervenciones. Una manera común de probar qué tan adecuada resulta la cirugía, en casos de difícil decisión, consiste en que el paciente pregunte a su médico que haría si se tratara de su familia. Si se hubiera realizado esta prueba, a más de la mitad de las mujeres se les habría extirpa-

do el útero cuando llegaran a una edad muy avanzada; esta proporción es mucho más alta que la que se considera de por sí elevada, del 35% de la población femenina en general²².

Se puede obtener una evaluación más completa de la atención médica y quirúrgica con los indicadores críticos de la atención, de manera que abarquen las listas más extensas de criterios explícitos para cada diagnóstico explícito, a las cuales ya he hecho mención.

El porcentaje de concordancia con dichos criterios, atribuyendo a los diferentes componentes pesos iguales o diferentes, puede tomarse como una medida sumaria de la calidad de la atención. El estudio de una muestra de casos hospitalarios de Hawai en el que se usó este método es particularmente notable, ya que proporciona una visión poco común de una parte importante de la atención a una población grande en su ambiente natural. El resultado del desempeño total fue del 71% de lo que se hubiera considerado como una concordancia perfecta de criterios. Por desgracia, no se ofrece una distribución de frecuencia de los resultados, y tampoco podemos juzgar si el 70% es un porcentaje bueno, malo o mediano. La aplicación del mismo método a una muestra concedidamente sesgada la atención de consultorio en Hawai, dio el funesto resultado de 41% con respecto a una concordancia total, valorada a partir de la información que aparece en las historias²³.

En mi opinión, un juicio decisivo de la calidad de la atención para cada caso no puede apoyarse únicamente en la concordancia con los criterios explícitos; debe basarse en una revisión de todos los hechos conocidos, por uno o más expertos, para llegar a emitir un juicio. El estudio de la calidad de la atención hospitalaria recibida por miembros de la Unión de Camioneros de la ciudad de Nueva York, constituye un ejemplo de esta tradición. Se les dio a dos médicos eminentes, por separado, el expediente completo de cada caso y se les pidió que valorasen la atención médica tomando como criterio la manera como hubieran tratado el caso. El resultado fue que opinaron que le 43% de los casos había recibido una atención médica no "óptima"²⁴.

Tanto en el estudio de Hawai como en el de los camioneros se intentó detectar los factores relacionados con la calidad de la atención. Si se toman ciertas libertades, se puede elaborar con ellos un cuadro compuesto²⁵. El principal factor relacionado con la calidad de la atención hospitalaria es la naturaleza del propio hospital. La mejor atención se da en los hospitales grandes, urbanos, afiliados a una universidad, y suele ser peor en los hospitales pequeños, ya sean urbanos o rurales. Otro factor es la especialización del médico: su influencia favorable es menor y se la advierte sólo cuando la práctica se limita a su área; pero toda vez que sale de esta puede resultar peor que un médico general. La importancia que otorga el hospital a la garantía de la calidad es mayor para el médico general, mientras que, fuera de los mejores hospitales, la especialidad del médico constituye la garantía más importante. Los médicos que trabajan en grupos grandes proporcionan la mejor atención hospitalaria; pero parece que esto se debe, sobre todo a que algunos son especialistas. En el caso de la atención ambulatoria, la práctica de grupo tiene una pequeña ventaja sobre la individual, pero los datos al respecto son más fidedignos. Quizá más importante que todas estas asociaciones sea la observación de que una gran parte de las variaciones en el desempeño aún no

ha podido ser explicada, lo cual indica que nuestras medidas pueden ser defectuosas que no comprendemos muchas cosas acerca de los factores determinantes del desempeño.

Estudios del resultado de la atención

La incidencia y la preevaluación de enfermedad e incapacidad, la incidencia de mortalidad y los niveles de longevidad son indicadores obvios de la salud de una población. Sin embargo, la atención médica hace sólo una aportación más bien pequeña entre los múltiples factores sociales y biológicos que determinan dichos resultados. Se necesita un refinamiento considerable para revelar los efectos de la calidad de la atención.

Los resultados pueden tomarse como medidas más precisas y específicas si se realiza una selección cuidadosa, de manera que correspondan con las categorías específicas de los pacientes, se puedan prevenir o lograr mediante una buena atención médica, y sean valorados únicamente después de haber corregido las características que influyen sobre el grado de éxito que se puede esperar, incluso de la mejor atención médica. Recientemente se ha ofrecido a manera de indicadores de calidad de la atención en las comunidades, una lista densa de medidas que se cree responden a la atención²⁶. También se ha sugerido que la etapa en la cual comienza a atender las enfermedades, o en la cual se admite por primera vez al paciente en el hospital, indicando algo sobre el grado de facilidad con que se logra el acceso a la atención y sobre la calidad de ésta²⁷. Para determinados diagnósticos y condiciones es también posible especificar los resultados cuya medición sería más útil, cuando medir esos resultados y cuales son las características del paciente que se deben de tomar en... a fin de aislar la aportación de la atención médica para lograrlos. Es mucho más difícil especificar el grado en que las variaciones en la calidad de la atención se reflejan en dichos resultados²⁸.

Se puede considerar que el análisis de la mortalidad y la morbilidad postoperatorias representa el tipo de es... sobre los resultados que tienen mayor especificidad y refina... Se sabe desde hace mucho tiempo que el índice de mortalidad postoperatorio varía considerablemente de un hospital a otro. Un ejemplo digno de atención es aquel en el que se observó una diferencia de 25 veces entre 34 centros médicos. Al corregir las diferencias entre los diversos centros médicos. Al corregir las diferencias entre los diversos centros médicos (en cuanto a factores como el tipo de operación, la edad y el estado físico del paciente), la amplitud se redujo a una diferencia de siete veces en algunas operaciones y de tres en otras²⁹. Estas grandes diferencias, carentes de explicación, resultaron a tal grado inquietantes que se realizó otro estudio en el que se hicieron todos los esfuerzos para corregir las características de los pacientes que pudieron haber ocasionado las diferencias observadas. Persistieron las diferencias reales y significativas, lo cual indica que las posibilidades de que se presenten complicaciones graves o muertes después de las mismas operaciones en pacientes con características similares pueden ascender al doble o al triple en algunos hospitales con respecto a otros³⁰. Sospecho que aun estas grandes diferencias no bastan para ilustrar todo el problema ya que no es seguro que en situaciones de alto riesgo los beneficios derivados de las operaciones superen siempre a los riesgos.

Cuando los resultados son usados para el seguimiento de la atención médica en una institución o programa, se requiere un análisis cuidadoso de todos los sucesos adversos de mayor importancia, a fin de mejorar las prácticas futuras³¹. Los médicos pueden ser más conscientes de las consecuencias de sus actos si se les convence de que especifiquen en forma anticipada que mejoría esperan en la salud de los pacientes de determinadas categorías, de manera que sus logros se puedan comparar con sus expectativas³². Pero ya sea que los resultados esperados se especifiquen o no con anticipación no hay forma de evadir la responsabilidad de examinar y evaluar la propia atención. Tales evaluaciones "retrospectivas" también pueden servir como una herramienta básica para la investigación. Los estudios tempranos sobre la mortalidad materna y neonatal de la Academia de Medicina de Nueva York, constituyen ejemplos notables. En el periodo de 1930-1932, se estimó con base a un cálculo "conservador", que el 66% de las muertes de mujeres durante el parto podría prevenirse, y el 61% de dichas defunciones se atribuyó a errores de criterio o técnica por parte del médico³³. En el periodo de 1950 a 1951 se estimó que el 42% de las muertes de recién nacidos no prematuros podía haberse prevenido y cerca del 80% de las muertes susceptibles de prevención se debía a errores de criterio y técnica del médico³⁴. En ambos estudios el tipo de hospital y las aptitudes del médico influyeron en grado considerable en los resultados, que equivalían a la vida misma. Estos hallazgos, profundamente inquietantes, dieron lugar a que se introdujeran una serie de controles, incluyendo revisiones de todas las muertes; a estos controles se ha atribuido parte de las mejoras registradas desde entonces. No obstante, un examen reciente de las tendencias de la mortalidad materna, en Michigan, de 1950 a 1970, indica que pese al descenso notable de la mortalidad, el porcentaje de muertes "susceptibles de prevención" ha aumentado de forma marcada; de alrededor del 60% a cerca del 80%³⁵. Cuanto más se elevan los estándares de atención parece más difícil alcanzar la perfección.

Combinación del proceso y el resultado

Existen dos métodos para evaluar la calidad de la atención que se pueden clasificar en una categoría aparte, ya que están diseñados para mostrar ciertos elementos de un sistema de atención, valiéndose de una combinación de mediciones del proceso y el resultado. El primero puede ser llamado el método de "trayectoria", que consiste en seleccionar una o más enfermedades o condiciones y en seguir a los pacientes, desde el momento en que solicitan la atención hasta cierto tiempo después de que esta se ha dado por terminada. De este modo es posible examinar los pasos en una secuencia que constituye, con demasiada frecuencia, una odisea de fallas acumuladas, y también documentar el efecto final de esta experiencia sobre la salud del paciente. En un estudio de este tipo, los iniciadores de este enfoque encontraron que, de un grupo de pacientes que llegaron a la sala de urgencias de un hospital urbano mostrando síntomas gastrointestinales, el 33% no se presentó a todos los exámenes recomendados; estos no se realizaron en una forma adecuada en el 12%, y en el 15% se encontraron alteraciones que no fueron debidamente tratadas. Con todo ello, el índice de

errores fue de un 60%. Cuando se tomaron en cuenta los efectos del tratamiento, se estimó que el encuentro del paciente con esta institución en particular tuvo un efecto saludable sólo en el 27% de los casos³⁶.

Si una parte de un mapa mental del sistema de atención médica que lo subdivide en campos de función y de responsabilidad, es posible seleccionar una cierta cantidad de diagnósticos o condiciones que sirvan como indicadores de la calidad de la atención en cada subdivisión. Todos los diagnósticos o condiciones funcionan a manera de "rastreadores" ("tracers"), y se puede considerar que el conjunto de rastreadores proporciona algo análogo a un conjunto de sondeos, de un terreno aún sin explorar³⁷.

Esta atractiva idea ha sido probada, en parte, utilizando como rastreadores la incidencia y el manejo de anemia, las infecciones del oído, la pérdida de la audición y los defectos visuales, para evaluar la atención médica de niños de 6 meses a 11 años de edad en determinadas zonas de Washington, D.C. De esta exploración se deriva una triste imagen, debida a que muchas de las patologías no han sido reconocidas, son susceptibles de prevención, o bien no han sido tratadas adecuadamente. Por ejemplo, el 12% de los niños entre 4 y los 11 años de edad necesita lentes, pero no los tiene; el 31% de los que tienen lentes no los necesita; el 37% no tiene la graduación adecuada, y en el 5% el uso de los lentes empeora la vista en lugar de mejorarla³⁸.

Evaluación de las estrategias de atención

La atención del paciente es una actividad planeada que involucra la acción de elementos específicos, de entre un conjunto potencialmente grande, y un ordenamiento adecuado, a fin de lograr determinados objetivos de diagnóstico y tratamiento. Se puede llamar estrategia a un plan de acción, así como al patrón de acciones que resultan de él. En mi opinión, la esencia de la calidad -esa cosa inasible que se llama "juicio clínico"- radica en la elección de la estrategia más adecuada para el manejo de cualquier situación dada. Las estrategias alternativas que un médico podría razonablemente elegir son susceptibles de especificarse en forma de un árbol de decisiones que indique las diferentes vías alternativas y sus respectivas consecuencias. Se puede asignar a cada una de ellas una probabilidad basada, de preferencia en hechos demostrados o, cuando esto no sea posible en opiniones autorizadas. El criterio para elegir la estrategia óptima para el paciente en cuestión consiste en el balance de beneficios, riesgos y costes esperados, evaluados conjuntamente por el médico y el paciente³⁹. La elaboración y el uso de modelos que incorporen los conocimientos existentes pueden ser muy útiles para llegar a una especificación más definitiva de la calidad ya que el mejor plan de acción sugerido por la intuición puede no ser el mejor de acuerdo con un análisis más formal de las decisiones. Es más, tales modelos, al revelar las deficiencias críticas en los conocimientos existentes, estimulan la investigación de manera que al final, se puede especificar firmemente cual es el manejo óptimo.

Los resultados de estos adelantos han comenzado a influir en el campo de la evaluación de la calidad. Quizás el primer paso ha sido la elaboración de "mapa de criterios" que sustituyen a las listas de criterios explícitos.

El "mapeo" es la representación de los pasos que se siguen para hacer un diagnóstico, detectar las complicaciones y seleccionar e implementar el tratamiento. En él se reconoce que existen formas alternativas aceptables para cubrir cada requerimiento (por ejemplo, el de contar con un diagnóstico válido) y también que las acciones sucesivas dependen de los hallazgos anteriores. Estos mapas de criterio se están usando para evaluar la calidad sobre una base experimental⁴⁰. El siguiente paso consistirá en conectar lo anterior con el enorme cúmulo de trabajo que se desarrolla, independientemente de las actividades de evaluación de la calidad, en el diseño y prueba de diversas estrategias de atención⁴¹. Desde luego, la comprobación empírica de la validez de dichas estrategias mediante minuciosas pruebas clínicas proporcionará el fundamento de toda evaluación de la calidad y, de hecho, la medicina clínica en su totalidad⁴².

El contexto de la monitorización en los EE.UU.

La idea de que el contenido de la práctica médica debe de someterse a una vigilancia constante ha ido cobrando fuerza y finalmente se ha constituido en un principio apoyado por la ley. El propósito ostensible es "garantizar la calidad", aunque tal vez esta meta sea demasiado ambiciosa pues lo más que se puede esperar es que constituya una "ayuda" o que produzca una "mejoría" en la calidad de la atención. Claro está que la calidad de la atención depende de muchos factores, incluyendo la calidad de los estudiantes, su educación, entrenamiento y socialización para convertirse en jóvenes profesionales; las oportunidades con que cuenten para continuar su educación y para renovarse; la disponibilidad de los recursos y el financiamiento que permitan la aplicación de todo el potencial de la ciencia médica, y los incentivos profesionales y económicos que influyen en el comportamiento de los médicos. La monitorización del trabajo del médico pretende funcionar como un incentivo más para que el desempeño sea adecuado.

Es posible afirmar que cuando los médicos trabajan juntos se lleva a cabo una monitorización informal, debido a que comparten los pacientes, se consultan formal o informalmente, realizan actividades de enseñanza, etc. Aunque la monitorización formal pueda ser la garantía de calidad menos importantes es, en definitiva, necesario. Es la única manera de obtener información fidedigna sobre el funcionamiento del sistema; resultado menos caprichoso y más justo que una información informal y parcial; puede ser un incentivo para la autocritica y el aprendizaje, y es una manera más en que la profesión médica puede demostrar su sentido de responsabilidad ante el público.

Tradicionalmente, tanto en los EE.UU. como en muchos otros países, los profesionales han sido en gran medida responsables de regular su propia conducta, con el interés de regular su propia conducta, con el interés de alcanzar normas más elevadas; en este contexto el gobierno se ha limitado a desempeñar un papel de apoyo y refuerzo. A este respecto, la medicina tiene, en términos generales, un registro de logros de los cuales puede sentirse orgullosa. Pero, en años recientes, ha aumentado la sensación de que debería esforzarse más, o bien renunciar a algunas de sus prerrogativas aceptando una supervisión externa. Muchos factores han contribuido a este estado de cosas. El más importante ha sido el cambio trascendental de un financiamiento indi-

vidual de la atención médica a uno colectivo por vía de los programas privados de seguros de salud. Durante muchos años, las compañías y organizaciones privadas de seguros de salud, así como los representantes de los grupos más grandes de compradores de seguros, se han mostrado descontentos con un aumento en los costes de la atención que no permite asegurar la necesidad ni la calidad de los servicios recibidos. Sin embargo, era poco lo que podían o querían hacer, más allá de objetar los abusos más obvios. Pero cuando el propio gobierno comenzó a realizar el mayor desembolso de todos, al instituir los programas *Medicaid* (Ayuda Médica) y *Medicare* (Atención Médica), se contó con los medios y eventualmente con la voluntad para convencerse de que el que paga manda. El enorme gasto de fondos federales constituyó, sin duda, el acicate más intenso para iniciar la acción; pero también había una preocupación por la calidad de la atención y una necesidad de asegurar la fiabilidad de los programas ante el Congreso Federal y de éste ante el electorado. Este último ahora está constituido por un público mejor informado y más exigente.

Hubo otros adelantos anteriores y paralelos a éstos. En primer lugar, la concentración gradual de una parte fundamental de la atención en el hospital, el cual surgió como el centro dominante de la práctica organizada. En segundo lugar, la creciente conciencia -por parte del público, de los miembros de los patronatos de los hospitales y de las cortes- de la responsabilidad del hospital sobre la atención proporcionada por sus médicos⁴³. En tercer lugar, el desarrollo de los métodos y la tecnología para la evaluación y la monitorización de la práctica real⁴⁴. Todos estos factores, en conjunto, prepararon el escenario y proporcionaron los medios y la oportunidad para que surgiera una iniciativa de ley audaz, que formaba parte de las enmiendas de 1972 al Acta de Seguridad Social⁴⁵.

De acuerdo con esta legislación, el país ha sido dividido en 194 áreas y se ha solicitado a los médicos que practican en cada una de ellas que formen, en el nivel local, Organizaciones para la Revisión de las Normas Profesionales (*Professional Standards Review Organizations* -PSRO-), las cuales tienen la responsabilidad de monitorear la aptitud y la calidad de la atención en hospitales, sanitarios y consultorios médicos. Estas organizaciones se encuentran en fases diversas de organización y operación. Aun es demasiado pronto para conocer que tan de efectivas serán, y para saber si las mejoras que producirán valen su enorme coste, esfuerzo y pérdida de libertad profesional.

En mi opinión, si las PSRO realizan a conciencia su cometido, es seguro que habrá una mejoría en la calidad, en el coste o en ambos⁴⁵. Sin embargo, si los médicos, como grupo, no pueden o no desean controlar su propia conducta, podrían surgir presiones para llevar a cabo una inspección más vigorosa por parte de agencias externas a la institución médica, incluyendo a las aseguradoras, los departamentos estatales de salud o una agencia del propio gobierno federal. De manera alternativa, puede concluirse que lo que se necesita es un cambio radical en el modo como están organizados los servicios y como se contrata y se paga a los médicos, a fin de fortalecer los incentivos para un comportamiento profesionalmente adecuado. Así, la confianza radicaría primariamente en la creación de las condiciones apropiadas para una buena práctica, más que en el temor de que la práctica insatisfactoria sea descubierta y reprobada. No obstante, aun en las mejores condiciones tendrá que mantenerse una

monitorización constante, ya que sin él la medicina no puede observarse a sí misma ni saber hacia donde se dirige.

Conclusiones

En el presente artículo se clasifican los principales enfoques en torno a la evaluación del procedimiento y del resultado de la atención médica, y se describen someramente algunos estudios ilustrativos y sus hallazgos. La necesidad manifiesta de garantizar y mejorar la calidad de la atención médica ha llenado a la institución de mecanismos que someten la atención médica a una constante revisión, a fin de detectar y corregir las deficiencias. En este artículo se examinan los acontecimientos que dieron lugar a la intervención del gobierno federal de los EE.UU. en estas actividades, mediante su auspicio a las PSRO. Se describen las características principales de las PSRO y se discuten sus posibles repercusiones. Aun no se puede saber como funcionarían, pero en caso de que no lograsen sus objetivos será difícil hacer frente a las presiones que conduzcan al planteamiento de soluciones más radicales.

Bibliografía y notas*

1. Payne BC. *Utilization Review: A Handbook for the Medical staff*. Chicago (EE.UU.): American Medical Association, 1965.
2. Jacobs CM, Cristoffel TH, Dixon N. *Measuring the Quality of Patient Care: The Rationale for Outcome Audit*. Ballinger, Cambridge (Massachusetts, EE.UU.), 1976;51.
3. Slee VN. *Ann Intern Med* 1974;81:97.
4. Donabedian A. *Milbank Mem Fund Q* 1966;44:166.
5. Sanzaro PJ, Williamson JW. *Medical Care* 1968;6:123. Starfield B. *Milbank Mem Fund Q* 1974;52:39.
6. Breslow L. *International Journal of Epidemiology* 1972;1:347.
7. Fanshel S, Bus JW. *Oper Res* 1970;18:1021. Berg RL, ed. *Health Status Indexes*. Hospital Researches and Educational Trust, Chicago, 1973. Ellinson J, ed. *International Journal of Health Services* 1976;6:377-538. Varios autores. *Health Services Research* 1976;11:332-528.
8. Paterson OL, Andrews LP, Spain RS, Greenberg BG. 2ª parte. *J Med Educ* 1956;31.
9. Clute KF. *The General Practitioner: A Study of Medical Education and Practice in Ontario and Nova Scotia*. Toronto. Canada: University of Toronto Press, 1963. Jungfer CC, Last JM. *Medical Care* 1964;2:71.
10. Janzen E. *Quality Nursing Care Assurance: Initial Survey*. Documento leído en la reunión anual de la APHA. Nueva Orleans, EE.UU., 23 de octubre de 1974.
11. Rosenfeld LS. *Amer J Public Health* 1957;47:856. Lyons TF, Payne BC. *Medical Care* 1974;12:714.
12. Ciocco A, Hunt H, Altman I. *Public Health Rep* 1950;66:21. Anderson H. *Amer J Public Health* 1969;59:275. Rosenberg S N, Gunston C, Berenson L, Klein A, *ibid.* 1976;66:21. Brook R H, Williams K. *Medical Care* 1976;14(Supl.).
13. Eislew CW, Slee VN, Hoffmann RG. *Ann Intern Med* 1956;44:144. Léase también la referencia 3.
14. Huntley RR, Steinhauser R, White KL, Williams TF, Martín DA, Pasternack BS. *J Chronic Dis* 1961;14:630.
15. Williamson JW, Alexander M, Miller GE. *J Am Med Assoc* 1967;201:938.
16. McDonald CJ. *Ann Intern Med* 1976;84:162. Bennett GO, Winnickoff R, Dorsey JL, Morgan M. *Assessing Physician Performance in Ambulatory Care*. San Francisco, EE.UU.: American Society of Internal Medicine, 1976;140-56.
17. McCarthy EG, Widmer GW. *N Engl J Med* 1974;291:1331.
18. McCarthy EG. *Mandatory and Voluntary Second Opinion Programs in the Greater New York Area With National Implications*. Conferencia Dorothy R. Eisenberg. Escuela de Medicina de Harvard, EE.UU., 1976.
19. Sparling JF. *Hospitals* 1962;36:62;56.
20. Lewis CE. *N Engl Med* 1969;281:880. Bunker JP, *ibid.* 1970;282:135. Wennberg J, Gittelsohn A. *Science* 1973;182:1002. Lichtner S, Pflanz N. *Medical Care* 1971;9:311. Vayda E. *N Engl J Med* 1973;289:1224.
21. Donabedian A. *Inquiry* 1969;6:4. Roemer MI, Schonick W. *Milbank Mem Fund Q* 1973;289:271.
22. Bunker JP, Wm B. Brown Jr. *N Engl J Med* 1974;290:1051.
23. Payne BC, Lyons TF, Dwarhius L, Kolton M, Morris W. *The quality of Medical Care: Evaluation and Improvement*. Chicago, EE.UU.: Hospital Research and Educational Trust, 1976.
24. Morehead MA, Donaldson R, *et al.* *A Study of the Quality of Hospital Care Secured by the Sample of Teamster Family Members in New York City*. Nueva York, EE.UU.: Columbia University School of Public Health, 1964.
25. Esto incluye información de un análisis más amplio de los datos de Payne, *et al.* En: Rhee S. *Medical Care* 1976;14:733.
26. Rutsein DD, Berenberg W, Chalmers TC, Child III CG, Fishman AP, Perrin EB. *N Engl J Med* 1976;294:582.
27. Gonnella JS, Goran MJ. *Medical Care* 1975;13:467. Gonnella JS, Louis DZ, McCord JJ. *Medical Care* 1976;14:13.
28. Brook RH, Davies-Avery A, Greenfield S, Harris LJ, Lelah T, Solomon NE, Ware Jr JE. *Medical Care* 1977;15(Supl.).
29. Moses LE, Mosteller F. *J Am Med Assoc* 1968;203:422.
30. Scott WR, Forrest Jr. WH, Brown Jr. WnB. *Organizational Research in Hospitals*. Shortell SM, Brown M, eds. Chicago, EE.UU.: Blue Cross Association, 1976;72-89.
31. SOSSUS Subcommittee on Quality of Surgery. *Surgery in the United States*. Chicago, EE.UU.: The American College of Surgeons and the American Surgical Association, 1976;2132-55.
32. Williamson JW. *J Am Med Assoc* 1971;218:564.
33. New York Academy of Medicine, Committee on Public Health Relations. *Maternal Mortality in New York City: A Study of All Puerperal Deaths, 1930-1932*. Nueva York, EE.UU.: Oxford University Press, Commonwealth Fund, 1933.
34. New York Academy of Medicine. Committee on Public Health Relations. *Perinatal Mortality in New York City: Responsible Factors*. Cambridge, Massachusetts: Harvard University Press, 1955.

*Las referencias bibliográficas se reproducen en su formato original con las notas del autor

35. Schaffner W, Federspiel CF, Fulton ML, Gilbert DG, Stevenson LB. *Amer J Public Health* 1977;57:821.
36. Brook RH, Stevenson RL. *N Engl J Med* 1970;283:904. Algunas de las cifras citadas difieren un poco de las que aparecen en el original, debido a que se ha realizado una recomputación a fin de obtener categorías mutuamente excluyentes.
37. Kessner DM, Kalk CE, Singer J. *N Engl J Med* 1973;288:189.
38. Kessner DM, Snow CK, Singer J. *Assessment of Medical Care for Children*. Washington D.C.: Institute of Medicine, 1974;174, 186,188.
39. McNeil BJ, Keeler E, Adelstein SJ. *N Engl J Med* 1975;293:211. Pliskin N, Taylor AK. *Costs, Risks, and Benefits of Surgery*. In: Bunker JP, Barnes BA, Mosteller F, eds. Nueva York, EE.UU.: Oxford University Press, 1977;5-27.
40. Greenfield S, Lewis CE, Kaplan SH, Davidson MB. *Ann Intern Med* 1975;83:761.
41. Véase por ejemplo: Ginsberg AS. *Decision Analysis in Clinical Patient Management with an Application to the Pleural Effusion Syndrome*. Santa Mónica, California, EE.UU.: The Rand Corporation, 1971. Schwartz WB, Gory BA, Kassirer JP, Essig A. *Am J Med* 1973;55:459. Mc Neil BJ, Adelstein SJ. *N Engl J Med* 1975;293:221. McNeil BJ, Hessel SJ, Branch WT, Bjork L, Adelstein SJ. *Journal of Nuclear Medicine* 1976;17:163.
- Tompkins RK, Burner DC, Cable WE. *Ann Intern Med* 1977;86: 481. También varios ejemplos en Bunker JP, et al., eds. Citado en la referencia 39.
42. Cochrane AL. *Effectiveness and Efficiency: Random Reflections on Health Service*. Londres: The Nuffield Provincial Hospital Trust, 1972.
43. Shain M, Southwick AE. *Public Health Rep* 1966;81:581. Curran WJ. *Amer J Public Health* 1971;61:177.
44. Donabedian A. *A Guide to Medical Care Administration. Volume II: Medical Care Appaisal-Quality and Utilization*. Nueva York, EE.UU.: American Public Health Association, 1969.
45. Congreso de los Estados Unidos, Reformas de Seguridad Social de 1972, Ley Pública 92-603, XCII Congreso, HRI., 30 de octubre de 1970. Ver: Professional Standards Review, Parte B, Título XI;101-17. Una excelente descripción de las primeras etapas de aplicación se encuentra en Goran MG, Kellogg JS, Fielding J, Jassee W. *Medical Care* 1875(suplemento de abril).
- Para una reciente evaluación de la efectividad de las PSRO, véase: Office of Policy, Planning and Research, Health Care Financing Administration, Department of Health, Education and Welfare, Professional Standards Review Organizations, 1978 Program Evaluation. Publicación HEW., núm. HCFA-0300, U.S. Government Printing Office, Washington DC., EE.UU., enero de 1979.

Calidad y Economía

Donabedian A, Wheeler JR, Wyszewianski L. Quality, cost and health: an integrative model. *Med Care* 1982;20:975-92

El artículo *Quality, Cost and Health: an Integrative Model* fue publicado en 1982 en *Medical Care*, pero a mi me había pasado desapercibido. Fue el propio Avedis quien me lo mostró durante el Congreso de la ISQua en Venecia el año 1995. El sabía de mi formación económica y me invitó a analizar su trabajo. Avedis no estaba muy satisfecho con este artículo, lo escribió en colaboración con economistas pero me dio la impresión que la colaboración no había sido muy fructífera. Ahora seis años más tarde y con motivo de este monográfico sobre Avedis Donabedian he vuelto a leer este artículo lo que me ha hecho recordar aquel congreso veneciano en el que estuve con Avedis.

El artículo que comento ha sido ignorado por el grupo de los economistas de la salud y también por los profesionales sanitarios dedicados a la calidad. El objetivo del mismo es muy ambicioso: integrar calidad y economía en un mismo modelo. El resultado es un modelo conceptual que puede resultar difícil de leer para los no economistas y difícil de entender para los economistas. Esto les pasa a las síntesis entre áreas de conocimiento alejadas, cuyo resultado no satisface a nadie. Sin embargo estos esfuerzos para integrar enfoques conceptuales son precisa y específicamente los que hacen avanzar la ciencia.

La economía de la salud no ha entendido nunca el mundo de la calidad, para Alan Williams calidad era igual a eficiencia, es decir hacer las cosas bien a la primera sin desperdiciar recursos. Una vez nos reunimos un grupo de economistas y de epidemiólogos para intentar poner en común nuestros lenguajes. Nos encallamos en las cuatro E: Eficiencia, Equidad, Eficacia y Efectividad. Las dos primeras son los paradigmas de la

Economía: producir eficientemente y repartir lo producido con equidad. Mientras que los dos segundos eran las preferidas de la Epidemiología: aplicar la medicina en una situación ideal o hacerlo en una situación real. El diálogo no fue posible, a cada grupo sólo le interesaba su tema.

El problema es: ¿se incluyen los costes en la definición de calidad? o bien ¿hay la calidad por un lado y los costes por otro?. Las definiciones tradicionales de calidad: técnica, servicio, satisfacción del paciente, etc. no incluyen el coste de la asistencia en la definición de calidad. El coste aparece más tarde, cuando el gerente o el político fija el límite presupuestario. Entonces el gestor se impone la tarea alcanzar el máximo de calidad con el dinero disponible. Avedis pretende incluir el coste en la definición de calidad técnica, y para hacerlo equipara calidad a salud. El artículo utiliza otras dos definiciones de calidad, la individual y la social. La primera incluye en la definición de calidad los costes y beneficios que recibe el paciente, mientras que la segunda incorpora los costes y beneficios sociales. Al incluir características personales y sociales, se distorsiona la definición técnica de calidad y las cosas son mucho más complejas.

El esfuerzo de Avedis de intentar sintetizar la calidad y la economía fue atrevido y merecía haber tenido continuidad, desgraciadamente todavía carecemos de un modelo que integre la calidad y la economía en la sanidad.

Lluís Bohigas

*Director General de Planificación Sanitaria.
Ministerio de Sanidad y Consumo*

Calidad, coste y salud: un modelo integrador*

Avedis Donabedian, John R.C. Wheeler, Leon Wyszewianski

En este artículo se presenta un modelo integrador de las relaciones entre el estado de salud, la calidad de la atención y la utilización de recursos. Se define la calidad de la atención médica en términos de resultados, medidos como las mejoras esperadas en el estado de salud que pueden atribuirse a la atención. El examen de la manera en que el uso ilimitado y eficiente de los recursos para la atención influye en la calidad así definida lleva a la definición absolutista de calidad. El análisis de la incidencia de los costes y beneficios individuales y externos de la atención sienta las bases de una distinción entre la definición individualizada de calidad, que depende de preferencias individuales, así como de la solvencia, y la definición social de calidad, que incluye consideraciones relativas a los beneficios externos, los costes sociales completos y las preferencias relativas a la distribución del bienestar. Se realiza otra distinción entre la eficiencia clínica y la eficiencia productiva. En el presente artículo se examinan en detalle las repercusiones de tres definiciones de calidad (absolutista, individualizada y social) y de los dos tipos de eficiencia (clínica y productiva) en la selección de estrategias óptimas de atención.

Gran parte de la preocupación actual en el campo de la atención médica se centra en el rápido aumento de los gastos en servicios de salud que se ha producido durante los últimos 20 años. Otro campo que ha atraído creciente atención durante los últimos diez años es la calidad de los servicios de salud. Estas dos inquietudes (el gasto en salud y la calidad de la atención) no son, por supuesto, independientes. A medida que los gastos aumentaban, el interés en la calidad de los servicios adquiridos también ha aumentado, no solo debido a la presión para determinar si esos servicios son, de hecho, de calidad satisfactoria, sino también debido a la mayor inquietud respecto de la magnitud de algunos gastos en relación con el incremento correspondiente de la calidad.

La adopción de las Enmiendas del Seguro Social en 1972, que estipulan el establecimiento de las Organizaciones de Revisión de Normas Profesionales (ORNP), fue una manifestación visible de la preocupación por la calidad de la atención frente al rápido aumento del costo de la atención médica. Por medio de esta ley, el gobierno federal, que ya en 1972 absorbía una parte importante de los gastos de atención médica, quería establecer un mecanismo para garantizar que se efectuaran pagos única-

mente por servicios que fuesen necesarios y de calidad satisfactoria, suministrados en el establecimiento menos costoso. La legislación relativa a las ORNP no puso en tela de juicio los conceptos o prácticas básicos prevalentes. Los métodos de las ORNP se basaron en mecanismos tradicionales de análisis por colegas y procedimientos de viejo cuño tales como el análisis de la utilización, los perfiles estadísticos y las auditorías médicas. La aplicación de estos métodos permaneció donde siempre había estado, es decir, en manos de los médicos y hospitales locales. Sin embargo, un aspecto fundamental es que la legislación sobre las ORNP no definió, y en realidad no pudo hacerlo, lo que constituía un servicio necesario de buena calidad, fuera de estipular que debía ceñirse a las prácticas y normas aceptables para la profesión.

Sin embargo, cuando se adoptó la legislación relativa a las PSRO, en otras esferas se estaban formulando preguntas más básicas sobre la utilidad de gran parte de la práctica médica. Cochrane¹ señaló varias prácticas de la medicina que son ampliamente aceptadas aunque no se ha comprobado su eficacia en forma concluyente. Posteriormente, Illich² fue aún más allá, argumentando que, en general, la atención médica es más perjudicial que beneficiosa, preconizando lo que Starr³, en una crítica convincente llama una nueva clase de "nihilismo terapéutico". En la obra de Fuchs aparecen dudas similares sobre la eficacia de la atención médica, aunque este autor hizo hincapié principalmente en las cuestiones de la eficiencia y la equidad de la asignación de recursos^{4,5}.

Basándose en la obra de Fuchs y Cochrane, entre otros, Havighurst y Blumstein examinaron en cierto detalle la amplia esfera de "los pros y los contras de la calidad y el costo de la atención médica"⁶. La relación entre coste y calidad ha sido examinada también desde un punto de vista más estrechamente arraigado en consideraciones sobre la calidad de la atención por Donabedian⁷⁻⁹ y Vuori¹⁰ y desde otro punto de vista por Phelps¹¹, quien propuso un método para realizar evaluaciones de coste-beneficio de los programas de control de calidad. Centrándose más específicamente en la atención de pacientes ambulatorios, Brook y Davies-Avery también han analizado los pros y los contras del costo y la calidad^{12,13}.

En este artículo presentamos un marco en el cual se examina sistemáticamente la relación entre la utilización de recursos y la calidad de la atención. El marco que proponemos inte-

gra conceptos sobresalientes de las obras mencionadas, así como otros conceptos pertinentes, y en particular aquellos relacionados con la evaluación del estado de salud.

Definición y evaluación del estado de salud

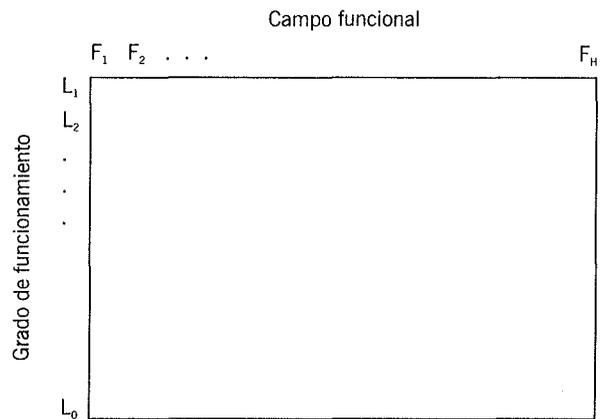
Comencemos con una definición de trabajo de la calidad de la atención médica en términos de resultados asociados a la atención recibida⁸. Específicamente, definimos la atención de la mejor calidad como aquella que produce la mayor mejoría prevista en el estado de salud, entendiendo la salud en términos generales en sus dimensiones física, fisiológica y psicológica. Si bien esta definición parece bastante sencilla, y tal vez hasta obvia, su explicación requiere un análisis de algunas relaciones complejas y distinciones sutiles. El concepto de calidad tiene dos componentes básicos: una definición y evaluación del estado de salud y de los cambios o diferencias en el estado de salud, y una definición de la atención médica vinculada a cualquier producto determinado expresado en forma de un estado de salud. En esta sección y en la siguiente abordaremos estas relaciones.

En una caracterización razonablemente completa del estado de salud se reconoce la índole pluridimensional de la salud, así como la relación entre la salud actual y el grado probabilístico de salud futura. Según Fanshel y Bush¹⁴, si evaluamos la salud en términos de grados mutuamente excluyentes de funcionamiento o actuación en distintos campos funcionales, la mejor manera de representar el estado de salud pluridimensional de una persona tal vez sea una matriz como la que se muestra en la Figura 1.

A efectos de este artículo conceptual, no necesitamos especificar los campos funcionales o los grados de funcionamiento o actuación. Simplemente a título de ejemplo podríamos aceptar, al igual que Breslow *et al.*, que la salud tiene "tres ejes: el físico, el mental y el bienestar social de la definición de la OMS"¹⁵. Después podríamos adoptar las escalas de actuación funcional que Breslow *et al.* prepararon para cada uno de esos campos funcionales^{16,17}. El "perfil del impacto de la enfermedad" es un ejemplo de una fórmula más compleja. Dicho perfil reconoce 14 campos del comportamiento, cada uno de los cuales se describe por medio de una escala de actuación o funcionamiento de múltiples posiciones¹⁸⁻²¹. También podríamos recurrir a la obra de Brook *et al.*, quienes idearon cuatro conjuntos de mediciones que corresponden a las cuatro dimensiones del estado de salud que señalaron: salud física, mental y social, y "percepciones generales sobre la salud"²². Este mismo grupo de investigadores ha abordado más recientemente los problemas de la combinación de las distintas escalas que representan la salud física, preconizando al mismo tiempo un modelo más parsimonioso que incluye solo los dos componentes principales del estado de salud: el físico y el mental^{23,24}.

Frente a una matriz de campos funcionales, por una parte, y de grados de funcionamiento o actuación por la otra, tenemos dos caminos. El primero consiste en asignar valores numéricos a cada una de las preferencias de los individuos o grupos en lo que atañe a los campos funcionales relacionados entre sí y a los grados de funcionamiento dentro de cada campo en relación con otros grados del mismo campo. De esta manera obtene-

Figura 1. Matriz del estado de salud



mos un valor numérico único para toda la matriz. Un aspecto intrínseco de este valor es la utilidad de cada estado de salud para un individuo o grupo²⁵.

El otro procedimiento consiste en reducir la matriz pluridimensional a una representación unidimensional, tal como preconizan Bush *et al.*, entre otros¹⁴. Una vez hecho esto, es posible usar diversos métodos para asignar un valor numérico a cada estado de manera tal que represente las preferencias (y, por lo tanto, la utilidad) de una persona o un grupo^{14,23-30}. De esta manera, la matriz original de estados de salud se puede resumir por medio de un vector S_t , cuyos elementos son indicadores de posibles niveles del estado de salud de una persona en el momento t :

$$\begin{bmatrix} S_{1t} \\ S_{2t} \\ \cdot \\ \cdot \\ S_{jt} \end{bmatrix}$$

En cualquiera de los dos procedimientos descritos, a la muerte, que es la ausencia de toda función, se asigna razonablemente un valor de 0. Al grado más alto de actuación que pueda definirse se asigna un valor más arbitrario, como 1 ó 100. Suponemos que el estado de salud se puede medir en una escala semilogarítmica, es decir, una escala caracterizada por intervalos iguales y un punto cero absoluto³¹. Esto significa que un $S_{jt}=20$ representa el doble del estado de salud que un $S_{jt}=10$, suposición que no es irracional. En nuestra opinión, la muerte representa un punto cero aceptable para la escala semilogarítmica. Si hubiese estados de salud menos deseables que la muerte, se les asignaría un valor negativo.

Existen indicios de que hay grupos de profesionales y consumidores que pueden indicar, de manera bastante confiable, los valores relativos de los distintos niveles del vector de esta-

dos de salud en una escala de intervalos aparentemente iguales, aunque los niveles en sí no sean necesariamente equidistantes uno del otro^{26,28}. Como es de suponer, las propiedades semilogarítmicas de nuestro indicador del estado de salud simplifican la tarea de obtener valores numéricos de los resultados de la atención que sean matemáticamente manipulables. Hay que reconocer también que la evaluación del estado de salud en una escala semilogarítmica basada en las preferencias de la gente en relación con distintos grados de funcionamiento significa que la utilidad total de la salud aumentará en una cifra constante con cada incremento en la medición del estado de salud. Es decir, la utilidad marginal del estado de salud permanece constante si la salud se mide en términos de preferencias.

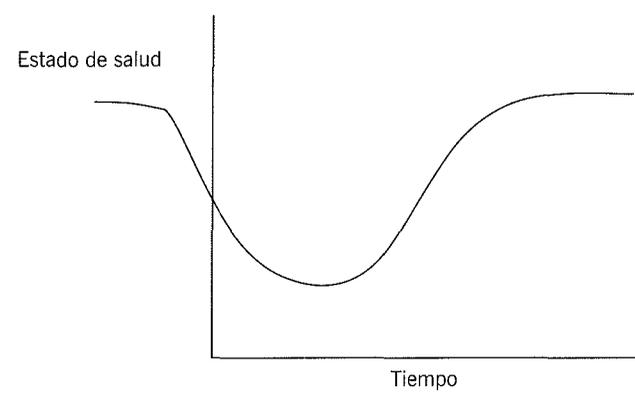
En nuestra evaluación del estado de salud de una persona incluimos la salud prevista para el futuro sumando al correr del tiempo los estados de salud previstos para el futuro, siendo la salud prevista para un año futuro la suma ponderada según la probabilidad de los posibles estados de salud correspondientes a ese año. Siendo S= suma de la salud actual y la salud prevista para el futuro, entonces:

$$S = \sum_t \sum_j p_{jt} S_{jt}$$

donde p_{jt} = la probabilidad de que la persona tenga el estado de salud j durante el período t . Para simplificar la notación y la exposición, por ahora dejaremos de lado la preferencia temporal por la salud suponiendo que los estados de salud actuales tengan los mismos valores que los estados de salud similares en un futuro más distante. En cuanto al período durante el cual se evalúa el estado de salud previsto, podría ser una vida definida conforme a las normas, como proponen Fanshel y Bush¹⁴, o la esperanza media de vida, como propone Sullivan³².

Somos conscientes de las numerosas dificultades, tanto conceptuales como operacionales, que surgirían al utilizar cualquiera de los dos procedimientos descritos. Reconocemos tam-

Figura 2. Posible curso de una enfermedad que se cura espontáneamente ante la ausencia de una intervención médica, que conduce a un estado de salud S_0 en un plazo determinado



bién que existen argumentos a favor del uso de un perfil de estados funcionales o de uno o más indicadores que representen directa o indirectamente el estado de salud en situaciones específicas, a menudo definidas desde el punto de vista clínico^{22,23}. No obstante, hemos elegido un indicador más completo e integrador del estado de salud porque es un componente necesario de una definición igualmente completa de calidad, la cual, en particular, incluye los aspectos técnicos de la calidad de la atención, así como las relaciones humanas que esta entraña⁹. Se pueden usar indicadores menos amplios del estado de salud, lo cual implicaría una definición incompleta de calidad, pero no alteraría de otra manera las propiedades del modelo que presentaremos. En todo caso, nuestro modelo demostrará la estrecha relación que existe entre la definición y evaluación de la salud, por una parte, y la definición y evaluación de la calidad, por la otra.

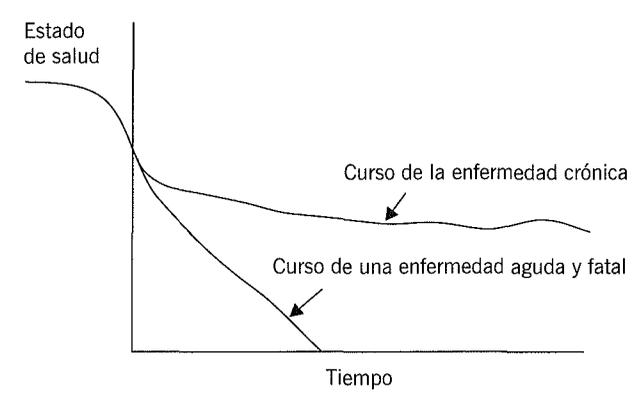
La relación entre atención médica y salud

La definición y la evaluación del estado de salud expuestas anteriormente pueden utilizarse para examinar la relación entre la atención médica recibida y el estado de salud. Veamos el caso de una persona que se ha enfermado. Si esta persona no recibe atención médica, la enfermedad seguirá su curso natural, que es probabilístico en su progresión y produce un estado de salud inesperado con el tiempo. Si S_0 representa la salud prevista con el tiempo sin intervención médica, entonces:

$$S_0 = \sum_t \sum_j P_{jto} S_{jt}$$

donde P_{jto} es la probabilidad de que la persona tenga el estado de salud j durante el período t si se adopta la estrategia o , es decir, si no se administra atención médica. Nótese que la probabilidad de encontrarse en cualquier estado de salud en el futuro depende de la estrategia médica adoptada.

Figura 3. Posibles cursos de enfermedades que no se curan espontáneamente ante la falta de intervención médica



Se pueden señalar varios tipos de cursos de la enfermedad³⁴. En un caso, la enfermedad se cura espontáneamente, en el sentido de que, a medida que pasa el tiempo, se puede prever que el estado de salud volverá a un nivel muy similar al existente antes del comienzo de la enfermedad, aunque no se reciba tratamiento. El curso previsto de este tipo de enfermedad y, por lo tanto, el nivel previsto de salud de la persona con el tiempo, tal vez tenga el aspecto de la Figura 2, donde tiempo = 0 corresponde al período inmediatamente posterior a la aparición de síntomas evidentes de la enfermedad y donde estado de salud = 0 indica la muerte.

Nótese que los puntos en el eje vertical corresponden a los elementos del vector S_t antes mencionado. Nótese también que S_0 , el valor previsto de la salud en ausencia de atención médica, se describe por medio del curso de la enfermedad en el diagrama, en el sentido de que S_0 , es la suma al correr del tiempo de los niveles de salud previstos en cada momento.

En otros casos, la enfermedad puede persistir y volverse crónica o conducir a la muerte; en ausencia de tratamiento médico, el paso del tiempo no lleva de vuelta a la persona al estado anterior a la enfermedad. La representación gráfica de estos casos puede tomar la forma de las curvas de la Figura 3. A efectos de la exposición que sigue, se tendrán en cuenta únicamente las enfermedades que se curan espontáneamente. Las relaciones establecidas podrán hacerse extensivas a otros casos sin dificultad.

En la mayoría de los casos, se puede contar con que la atención médica alterará el curso de los estados de salud previstos después del comienzo de la enfermedad. La magnitud de la alteración y si dicha alteración consiste en una mejoría o en un empeoramiento dependen de la pericia de quienes suministran los servicios de salud, la disponibilidad de recursos, el grado de desarrollo de la ciencia y la tecnología de la medicina, y varios factores que influyen en la gravedad de la enfermedad y en la capacidad de recuperación del paciente. En el mejor de los casos, el enfermo será manejado por un experto o médico "ideal" que trabaje en un medio donde los recursos no estén limitados*. Por manejo entendemos una estrategia de atención, y suponemos que una estrategia se puede especificar en términos de la clase y la cantidad de servicios proporcionados, el orden y la oportunidad de esos servicios, y las normas de decisión que llevan a optar por ciertas maneras de proceder entre muchas otras posibles. Por médico ideal entendemos el que selecciona y lleva a cabo la estrategia de atención que permite mejorar al máximo el estado de salud sin derrochar recursos. Además, adoptamos la suposición simplificadora de que las estrategias de atención seleccionadas por el médico ideal, así como aquellas seleccionadas por los médicos "no ideales" que examinaremos a continuación, se llevan a cabo con un grado de eficiencia óptimo. Es decir, no se derrochan recursos porque no existe una subutilización de instalaciones, las tareas no se realizan en forma incorrecta, la coordinación no es insuficiente, ni existen otros factores de ese tipo. Más adelante en este artículo analizaremos las repercusiones de esta suposición.

*A fin de aclarar la exposición, en este artículo utilizamos al médico como prototipo del profesional independiente que decide qué tratamiento recibirá el paciente. Nuestro análisis se aplica igualmente a cualquier otro profesional que tome decisiones análogas sobre el tratamiento en general.

El estado de salud previsto, sumado al correr del tiempo, relacionado con el manejo del caso por el médico ideal (I) y sin limitaciones (U) S_{IU} se expresa por medio de la fórmula siguiente:

$$S_{IU} = \sum_t \sum_j P_{jIU} S_{jt}$$

donde P_{jIU} es la probabilidad de que la persona tenga el estado de salud j en el momento t si el caso es manejado por el médico ideal, cuyas decisiones con respecto al tratamiento no están restringidas. La progresión de los estados de salud en una enfermedad que se cura espontáneamente sujeta a este tratamiento ideal y sin restricciones podría expresarse gráficamente como la curva superior de la Figura 4.

La separación entre las dos progresiones de los estados de salud que se muestran en la Figura 4, que puede observarse integrando las expresiones correspondientes a S_{IU} y S_0 sobre t y averiguando la diferencia, representa la mejoría prevista del estado de salud resultante de la atención proporcionada por el médico ideal y sin restricciones. Supongamos que Δ_{IU} representa esta mejoría esperada, donde $\Delta_{IU} = S_{IU} - S_0$. Utilizando la terminología más tradicional del control de la calidad, este es el resultado, en términos de salud, de la atención proporcionada por el médico ideal y sin restricciones.

Ahora bien, si limitamos el valor en dólares de los recursos de que dispone nuestro médico ideal a R_{IC} , sin restringir los tipos de recursos que puedan utilizarse, y le pedimos al médico que maneje la atención del mismo paciente, el estado de salud previsto, sumado al correr del tiempo y expresado por medio de la fórmula siguiente:

$$S_{IC1} = \sum_t \sum_j P_{jIC1} S_{jt}$$

Figura 4. Posible progresión de los estados de salud en relación con una enfermedad que se cura espontáneamente en los casos en que no se administra tratamiento y en los casos en que es tratada por un médico ideal sin limitaciones en cuanto a la disponibilidad de recursos

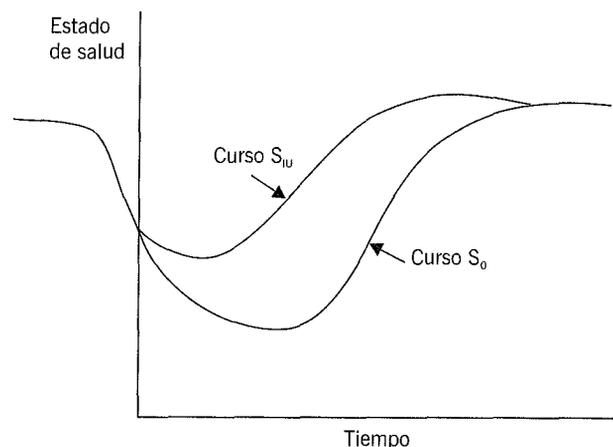
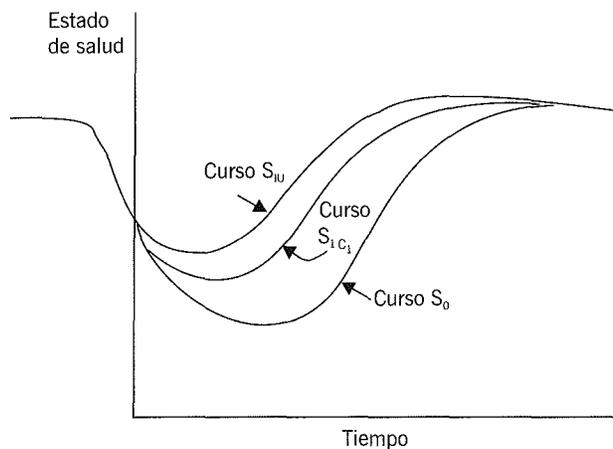


Figura 5. Posible progresión de los estados de salud en relación con una enfermedad que se cura espontáneamente en los casos en que no se administra tratamiento y en los casos en que es tratada por un médico ideal con y sin limitaciones en cuanto a la disponibilidad de recursos



será un poco más bajo. En el caso de una enfermedad que se cure espontáneamente, la recuperación tal vez sea más lenta, tal como indica la curva central de la Figura 5.

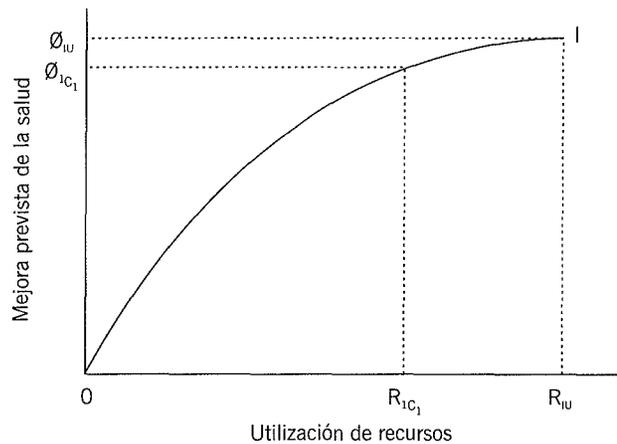
La separación entre el curso de la enfermedad que lleva a S_{IC_1} y el curso que lleva a S_0 representa la mejoría prevista, ϕ_{IC_1} en este primer caso de limitación de los recursos, mientras que la separación entre el curso de la enfermedad que conduce a S_{IC_1} y el que conduce a S_U representa la disminución total que puede atribuirse a este grado de limitación de los recursos.

Si los recursos de que dispone el médico ideal se restringen gradualmente, obtenemos una serie de puntos que establecen una relación entre el producto previsto en términos de salud (O) y la utilización de recursos. Aunque no sabemos qué aspecto tendrá la curva descrita por estos puntos, tal vez tenga la forma parabólica general de OI de la Figura 6.

Por supuesto, existen otras formas posibles, como la clásica forma de S o una forma exponencial. De hecho, debido a la índole indivisible de la atención médica, la relación tal vez aparezca como una función gradual en una gama de utilización de recursos. Para facilitar la exposición, suponemos que es posible aproximarse a la relación entre utilización de recursos y mejoría prevista mediante una curva parabólica continua que indique que la mejoría marginal de la salud disminuye en forma constante a medida que aumenta la utilización de recursos. La exposición que sigue no depende de esta suposición simplificada; nuestra hipótesis puede aplicarse fácilmente a las curvas que tengan una forma diferente de la OI de la Figura 6.

La curva OI de la Figura 6 muestra la relación entre la atención médica y el producto en términos de salud cuando el tratamiento es manejado por el médico ideal. Por supuesto, el médico ideal es algo así como un mito creado para definir el límite superior de los resultados que podrían obtenerse con re-

Figura 6. Relación hipotética entre la utilización de recursos y la mejora prevista de la salud correspondiente a las estrategias de atención seleccionadas por el médico ideal



ursos específicos. En la vida diaria, la mayoría de los puntos que corresponden a la relación observada entre el estado de salud previsto y la utilización de recursos están por debajo y a la derecha de la OI, porque la mayoría de los episodios de enfermedad no son manejados por el médico ideal.

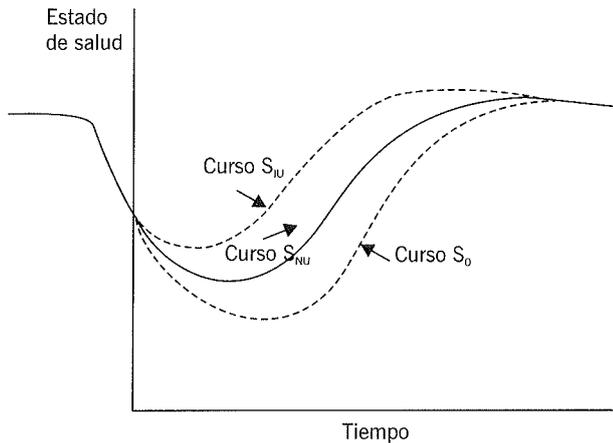
Para ver la manera en que podrían producirse desviaciones respecto de la OI, consideremos qué ocurriría si se pidiera a un médico típico o "no ideal" que atendiera a nuestro paciente hipotético. Supongamos primero que el médico no ideal (N) dispone de recursos ilimitados (U). El resultado previsto en términos de la salud del paciente en esta situación se expresa por medio de la fórmula:

$$S_{NU} = \sum_t \sum_j P_{j|NU} S_{jt}$$

donde $P_{j|NU}$ es la probabilidad de que la persona tenga el estado de salud j en el momento t si su atención está a cargo de un médico no ideal cuyas decisiones no están sujetas a restricciones. En el caso de una enfermedad que se cure espontáneamente, la progresión de los estados de salud tal vez tome la forma de la Figura 7.

Evidentemente, el médico no ideal obtiene un incremento más pequeño de la salud sumada al correr del tiempo que el médico ideal cuando no se colocan limitaciones a los recursos utilizados por ninguno de ellos. Además, se puede postular que, cuando se limitan los recursos, el médico no ideal obtendrá por cada dólar gastado un producto más pequeño en términos de salud que el médico ideal. En cambio, el médico no ideal utilizará más recursos que el médico ideal para obtener cualquier resultado específico en términos de salud. Estas relaciones entre la utilización de recursos y los resultados en términos de salud se muestran en la Figura 8.

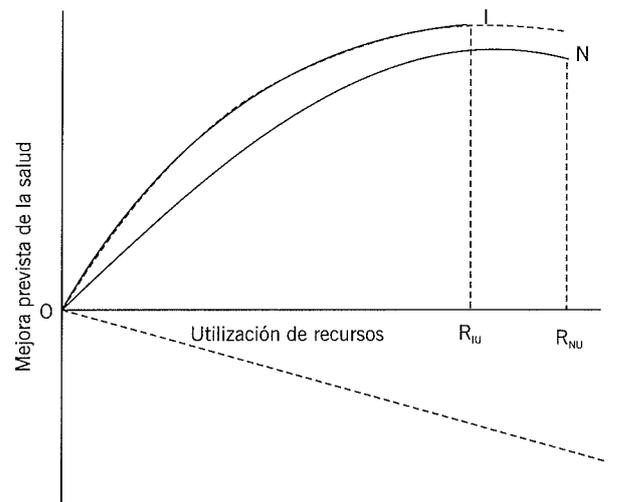
Figura 7. Posible progresión de los estados de salud en relación con un enfermedad que se cura espontáneamente en los casos en que no se administra tratamiento y en los casos en que es tratada por un médico ideal y por un médico no ideal sin limitaciones en cuanto a la disponibilidad de recursos



En la Figura 8, la curva OI muestra la actuación del médico ideal. La curva ON muestra la actuación del médico no ideal típico o corriente en una localidad determinada. Además de presentar las relaciones descritas, una comparación de las dos curvas permite realizar otra distinción importante. El médico ideal no utilizará recursos que excedan de los necesarios para conseguir los mayores beneficios posibles para la salud. En cambio, el médico no ideal corriente se deja llevar por la disponibilidad excesiva de recursos y emplea estrategias de atención que presentan mayores riesgos para la salud del paciente. Por esa razón, la curva ON en realidad comienza a descender después de alcanzar un nivel determinado de utilización de recursos.

Huelga decir que en cualquier localidad hay un número relativamente grande de médicos que no son ideales y que difieren considerablemente en cuanto a sus conocimientos, criterio y pericia. Los resultados alcanzados por esta población de médicos en el tratamiento de cualquier enfermedad específica en un grupo de pacientes puede expresarse por medio de un gran conjunto de puntos que caerán, con distintos grados de densidad, en el espacio limitado por las líneas de puntos de la Figura 8. Como ya se dijo, todos estos puntos caen debajo y a la derecha de la OI. Además, tomamos en cuenta los casos de atención perjudicial en todos los niveles de utilización de recursos, aunque postulamos que tanto la magnitud como la probabilidad del perjuicio son mayores cuando los recursos utilizados son excesivos. Aunque esperamos que esta situación sea rara, es posible imaginar a un médico que habitualmente haga más daño que bien a sus pacientes con cualquier nivel de utilización de recursos. Lo más cercano a la actuación de este médico sería la línea de puntos más baja de la Figura 8. En cambio, la actuación del mejor médico de la localidad tal vez se acerque a la del ideal hipotético. Sin embargo, en términos más estrictos

Figura 8. Relación hipotética entre la utilización de recursos y la mejora prevista de la salud correspondiente a las estrategias de atención seleccionadas por el médico ideal y por el médico no ideal



las dos líneas de puntos de la Figura 8 no representan la actuación del mejor y del peor médico de una localidad, sino que constituyen los límites superior e inferior de los resultados terapéuticos obtenidos por todos los médicos de la localidad que atienden a todos los pacientes con una enfermedad determinada. La actuación del médico "corriente", indicada por la curva ON en la Figura 8, también es un concepto estadístico que representa la actuación media, mediana o modal de todos los médicos de la localidad que atienden a todos los pacientes de un tipo determinado. La actuación de un médico en particular, especialmente cuando se la observe durante un período relativamente breve, probablemente sea mucho más errática y la curva de actuación corriente correspondiente a ese médico sería difícil de especificar por anticipado.

La distinción que hemos hecho entre la atención proporcionada por el médico ideal y la suministrada por el médico no ideal tiene ciertas similitudes con la distinción entre atención "eficaz" y "efectiva". Según la definición de Williamson, eficacia se refiere a "la medida en que pueda demostrarse que una intervención médica es beneficiosa en condiciones óptimas de atención clínica", mientras que la efectividad se refiere a "la medida en que realmente se obtengan en la práctica los beneficios que podrían alcanzarse en condiciones clínicas óptimas"³⁵.

El coste y la definición de la calidad

Hemos dicho que el punto I de la curva OI de las Figuras 6 y 8 indica la mayor mejoría posible del estado de salud que se podría conseguir con la intervención médica, teniendo en cuenta los conocimientos y las técnicas médicas actuales. También dijimos antes que, al principio, definiríamos la calidad óptima de la

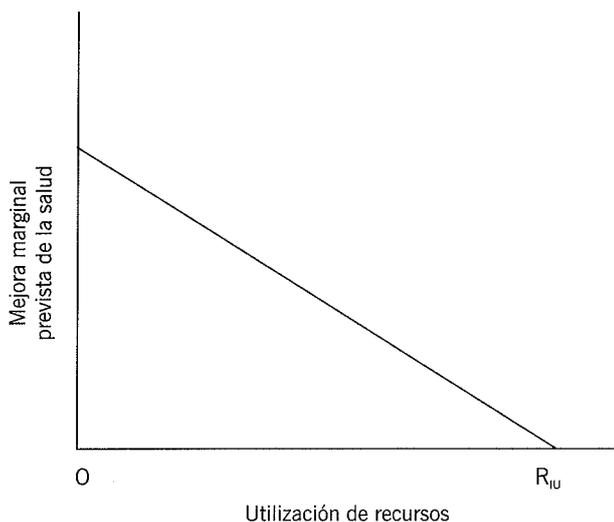
atención como aquella que permite mejorar al máximo el estado de salud. Por lo tanto, sobre la base de esa definición, la estrategia de atención empleada para producir O_{IU} a un costo de recursos de R_{IU} representa la calidad óptima de la atención médica. Este concepto de calidad es puramente médico en el sentido de que refleja lo máximo que se puede hacer para el paciente en el marco de los conocimientos técnicos actuales. Para una conceptualización más completa de la calidad se necesita considerar la relación entre el valor de la mejoría del estado de salud y los recursos necesarios para producir dicha mejoría.

La definición de estado de salud empleada en el presente artículo refleja un conjunto de preferencias por un nivel de funcionamiento en comparación con otro, pero no refleja preferencias por la salud en comparación con preferencias por otros productos y actividades valiosos. La curva OI de la Figura 8 puede considerarse como una curva del producto total en relación con los recursos médicos utilizados para mejorar la salud. La Figura 9 muestra el producto marginal de manera análoga a la relación del producto total, siendo el producto marginal definido como el cambio en la mejoría prevista asociada al dólar adicional gastado.

Matemáticamente, la mejoría marginal prevista, o O marginal, es la pendiente de la curva de la mejoría total prevista. Por consiguiente, en la cima de la curva OI de la Figura 8, donde la utilización de recursos es R_{IU} , la mejoría marginal prevista relacionada con la utilización de recursos adicionales es cero en la Figura 9.

El valor para una persona enferma de la mejoría marginal del estado de salud previsto indicado por el precio que la persona está dispuesta a pagar por esa mejoría, depende de la relación entre la utilidad de la salud y la utilidad de otros productos y actividades para la persona. Específicamente, el precio que

Figura 9. Relación entre la utilización de recursos y la mejoría marginal prevista de la salud



una persona está dispuesta a pagar por una mejoría marginal del estado de salud previsto, es decir, su valor en dólares, se expresa por medio de la relación:

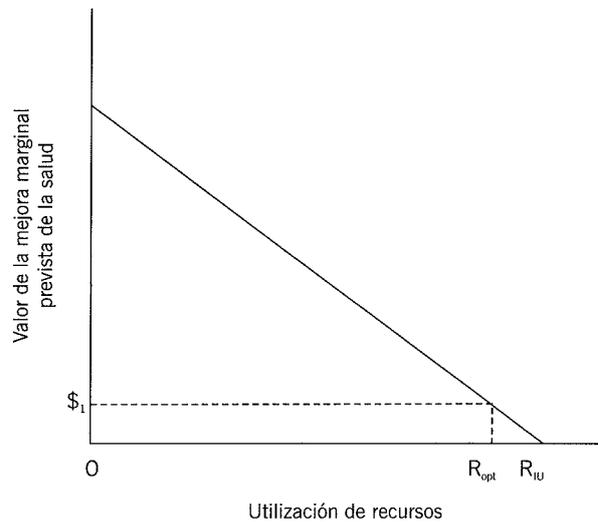
$$\text{Valor} = \frac{\text{Utilidad marginal de la mejoría prevista de la salud}}{\text{Utilidad marginal del dinero}}$$

La utilidad marginal del dinero representa la utilidad de otros productos y actividades que pueden comprarse con dinero, y se supone que es una función decreciente de la riqueza del individuo. Se parte de la hipótesis de que la utilidad marginal de la mejoría de la salud prevista es constante, sobre la base de la definición y evaluación del estado de salud presentadas antes en este artículo. Específicamente al definir la salud en términos de las preferencias de la gente en relación con ciertos grados de funcionamiento, podemos especificar una función de utilidad lineal para la salud, lo cual nos da una utilidad marginal constante de la salud**. En consecuencia, el valor que una persona asigna a la mejoría marginal de la salud prevista depende de la riqueza de la persona y de su preferencia por la salud en comparación con otros productos básicos y actividades***. Por consiguiente, la relación entre el valor de una mejoría marginal del estado de salud previsto y los recursos utilizados tiene la forma general que se presenta en la Figura 10.

**Si la salud se evalúa de otra manera, independientemente de la preferencia de la gente, en general la función de la utilidad de la salud será no lineal. En este caso se invocará la hipótesis más común de la utilidad marginal decreciente de la salud. El análisis subsiguiente mostraría los mismos resultados, aunque las representaciones gráficas serían más complejas.

***En aras de la simplicidad dejamos de lado la aversión a los riesgos. Su inclusión no influiría en las relaciones establecidas en esta parte del artículo.

Figura 10. Utilización de recursos, valor de la mejoría marginal prevista de la salud y estrategia óptima



La pendiente de la curva es más empinada para las personas más adineradas y para las que asignan un valor relativamente alto a la salud, y más gradual para las personas de más bajos ingresos y para las que asignan un valor relativamente más bajo a la salud. El eje vertical de la Figura 10 se define con mayor exactitud como el valor en dólares del aumento adicional del estado de salud previsto que resulta de un incremento de un dólar en los gastos. El coste del gasto de un dólar adicional es, por supuesto, un dólar. Por lo tanto, los costos y beneficios marginales del tratamiento médico se equilibran con una utilización de recursos de R_{opt} de manera tal que la utilización de recursos adicionales para el tratamiento beneficiará al individuo menos que gastar esos recursos adicionales en otro producto o actividad. Desde el punto de vista del bienestar individual, en contraposición al estado de salud individual, la estrategia óptima de atención resulta en un gasto de R_{opt} , un nivel más bajo de utilización de recursos que R_{IU} . Por consiguiente, la máxima calidad de la atención médica, en términos de su efecto en el bienestar individual, es la estrategia de atención vinculada a una utilización de recursos de R_{opt} .

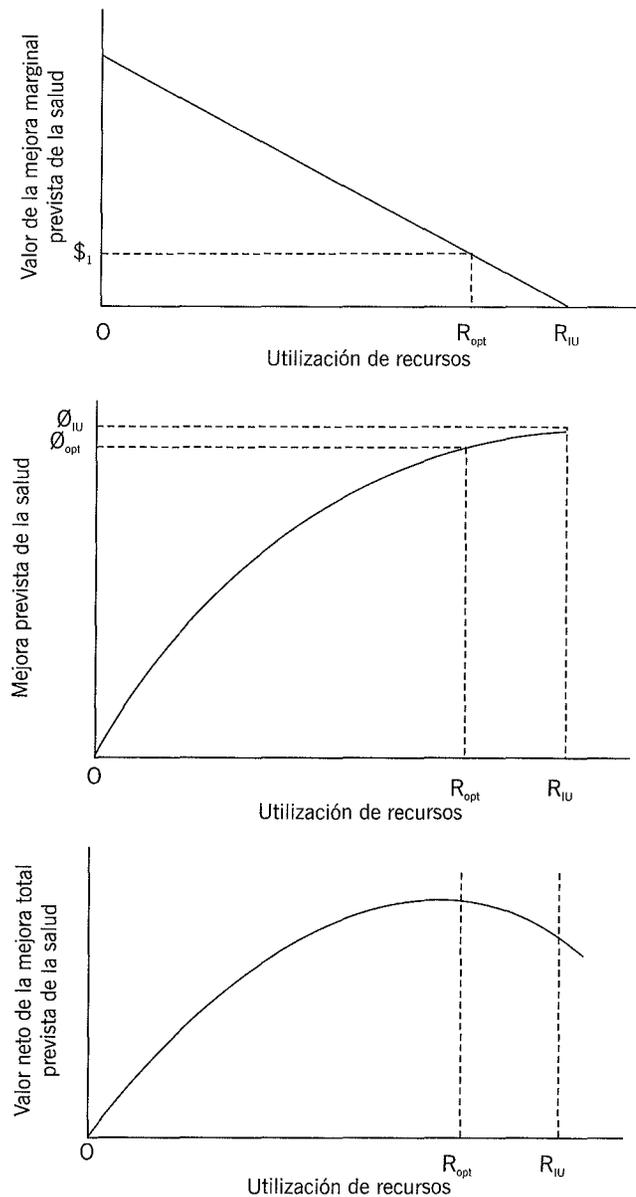
Cabe destacar que esta definición de atención óptima depende de la relación entre la utilización de recursos y el resultado cuando el tratamiento es manejado por el médico ideal. Como esta relación difiere de un médico a otro, la estrategia óptima de atención también variará, pero el nivel óptimo para el médico ideal sigue siendo la norma para la comparación.

En la Figura 11 se combina la figura anterior con otras dos figuras para mostrar el efecto de la utilización del nivel de recursos que corresponde a R_{opt} tanto en la mejoría prevista del estado de salud como en el valor neto de la mejoría prevista. Este valor se determina multiplicando la mejoría prevista de la salud (\emptyset) por el valor por unidad de mejoría, expresado en la fórmula precedente, y restando de este producto los recursos utilizados. Nótese que, aunque con una utilización de recursos de R_{opt} se puede aumentar un poco la mejoría prevista del estado de salud gastando recursos adicionales, el valor de ese aumento es muy pequeño. De hecho, el valor de la mejoría prevista del estado de salud es inferior al valor de los recursos que habría que utilizar para producir la mejoría, de manera tal que el valor neto de la mejoría total prevista del estado de salud disminuye. Por lo tanto, no vale la pena realizar el gasto adicional. Desde este punto de vista más global, una estrategia de atención que produzca una mejoría prevista del estado de salud de \emptyset_{IU} a un costo de R_{IU} en términos de recursos es de calidad inferior a una estrategia que produzca una mejoría prevista del estado de salud de \emptyset_{opt} a un costo de R_{opt} .

La calidad, el coste y la eficiencia de la producción

Hasta ahora, al preparar nuestro modelo de la relación entre calidad y coste nos hemos basado en la suposición de que el coste varía según la estrategia de atención empleada por el médico. El médico ideal siempre combina los servicios, controla su duración y los ordena de la manera más eficiente que sea posible para producir el mayor incremento de la salud con una cantidad específica disponible o permisible de gastos. Otros médicos emplean estrategias menos eficientes, y entonces producen menos salud por cada dólar gastado en la atención.

Figura 11. Efecto del empleo de la estrategia óptima en la mejora prevista de la salud y en el valor neto de la mejora total prevista de la salud

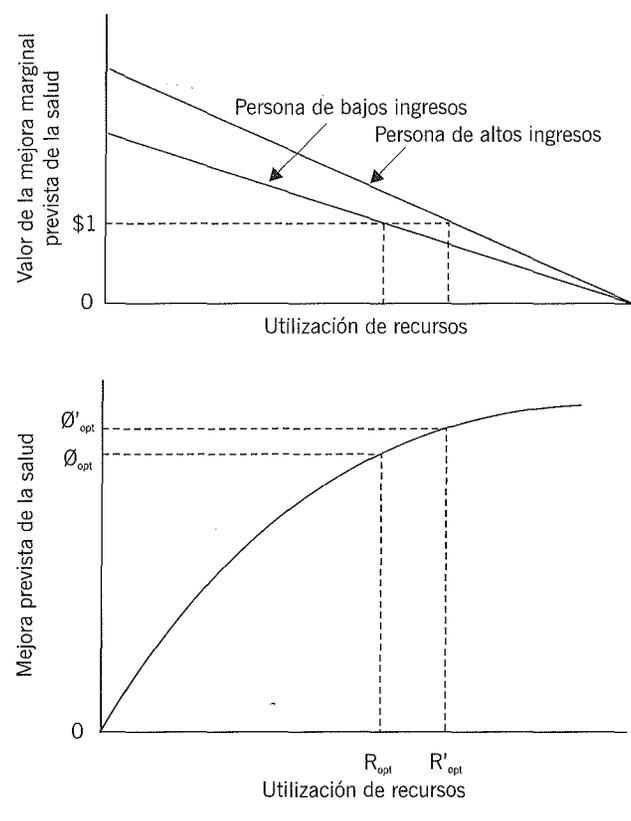


La eficiencia de las estrategias de atención puede llamarse "eficiencia clínica". Existe también otra clase de eficiencia, que guarda relación con la manera en la cual se producen los servicios y por lo general se llama "eficiencia productiva". Por ejemplo, la atención no se produce de manera eficiente si el hospital está medio vacío, si el laboratorio no informa rápidamente sobre un resultado o si se usan profesionales altamente especializados para hacer tareas que otros empleados menos

especializados podrían realizar igualmente bien a un coste más bajo. De ello se deduce que es posible obtener mayores incrementos de salud (o más calidad) por dólar gastado: 1. combinando los servicios, controlando su duración y ordenándolos en estrategias más eficientes; es decir, aumentando la eficiencia clínica, y 2. produciendo servicios en forma más eficiente; en otras palabras, aumentando la eficiencia productiva.

Si bien la eficiencia clínica es un componente fundamental de nuestra definición de calidad, preferimos excluir la eficiencia productiva de nuestra definición de calidad porque no entraña el uso de un juicio clínico. Sin embargo, reconocemos la importancia decisiva de la eficiencia productiva para la organización y el suministro de servicios de salud. Las mejoras de la eficiencia productiva nos permitirán alcanzar el nivel actual de calidad a un costo más bajo. Otra opción consiste en producir una mayor cantidad de servicios en los cuales la combinación de calidades permanezca como hasta ahora. Para mejorar la calidad aun más se necesitaría un cambio en las estrategias de atención. En consecuencia, aunque la eficiencia productiva es un componente de la calidad del sistema que produce atención, no es un componente de la calidad de la atención en sí.

Figura 12. Efecto de los ingresos en la estrategia óptima de atención



La calidad, los ingresos y las preferencias por la salud

Ya hemos examinado las repercusiones de la definición de calidad en términos de la capacidad y voluntad del paciente para realizar gastos con el fin de mejorar el resultado previsto en términos de salud. Es decir, los ingresos del paciente y sus preferencias por la salud influyen en la especificación de la estrategia óptima de atención. Si las preferencias son estables, un nivel de ingresos más bajo entraña una menor voluntad de pagar por la mejoría del estado de salud previsto. Por consiguiente, la curva que indica el valor de la mejoría marginal del estado de salud previsto para una persona de bajos ingresos está por debajo de la curva correspondiente a una persona de ingresos elevados y tiene una pendiente menos marcada, tal como se observa en la Figura 12. Por lo tanto, si lo único que se desea es alimentar al máximo la utilidad para la persona, la estrategia óptima de atención para una persona de bajos ingresos consiste en un gasto menor (R_{opt}) y una mejoría menor del estado de salud previsto (\emptyset_{opt}) que en el caso de la estrategia óptima para una persona de ingresos elevados. En otras palabras, este análisis lleva a la conclusión inquietante e incluso inaceptable de que la definición de calidad óptima depende de la capacidad de una persona para comprar servicios.

El mismo razonamiento se aplica cuando las preferencias relativas por la salud varían y los ingresos permanecen constantes. Si la persona A prefiere los bienes materiales a la salud relativamente más que la persona B, la estrategia óptima consistiría en destinar menos recursos a la mejora de la salud de A que de B. Una vez más, la definición de calidad está arraigada en el valor que la persona asigne a la salud. En la práctica, por supuesto, el valor del dinero o el valor de la salud varían simultáneamente y dependen de muchos factores, como los ingresos, la educación y la ocupación. En consecuencia, si se adopta esta perspectiva, la estrategia óptima de atención o, por lo tanto, la definición de calidad, variarán de una persona a otra, según las preferencias del individuo.

Además de las preferencias relativas por la salud y el dinero en un momento dado, la variabilidad de las preferencias individuales por la salud y el dinero con el tiempo también está comprendida en la especificación de la estrategia óptima. Específicamente, la actualización de los beneficios de salud futuros a una tasa elevada a fin de reflejar una preferencia importante por el bienestar actual en comparación con el bienestar futuro conduce a una reducción del valor de la mejoría marginal prevista del estado de salud y, por lo tanto, disminuye R_{opt} y \emptyset_{opt} . Un aspecto más importante tal vez sea la influencia de las preferencias temporales en la selección de estrategias que difieren en cuanto al momento en que presentarán riesgos y a los beneficios previstos. McNeil *et al.* han presentado un ejemplo instructivo del efecto de dichas preferencias en la opción entre el tratamiento médico y quirúrgico para el cáncer de pulmón³⁶.

De ello se desprende que hay dos clases de definición de calidad. La primera, que se acaba de abordar, corresponde a lo que Donabedian ha llamado la definición "individualizada"⁹. Se basa en la suposición de que el individuo paga todos los costos y recibe todos los beneficios, y refleja el valor que cada individuo asigna a la salud y a otros bienes, así como las preferencias

temporales de cada individuo. Este enfoque de la definición de la calidad implica que no existe una estrategia que sea mejor que todas las demás, sino que hay una estrategia que es la mejor para cada persona en cada situación. Se contraponen a la definición de calidad que hace caso omiso del coste, descrita antes en este artículo. Vuori¹⁰ la ha llamado calidad "absoluta", en tanto que Donabedian⁹ la ha llamado definición "absolutista". Estos calificativos son particularmente apropiados si el médico realiza una valoración independiente de la conveniencia de distintos resultados en términos de salud sin consultar con el paciente. En estas circunstancias, por lo menos en teoría quizá sea posible especificar la mejor estrategia, tal como parecen hacer a veces los criterios y normas de control de calidad.

La estrategia óptima para el individuo y la estrategia óptima para la sociedad en la definición de calidad

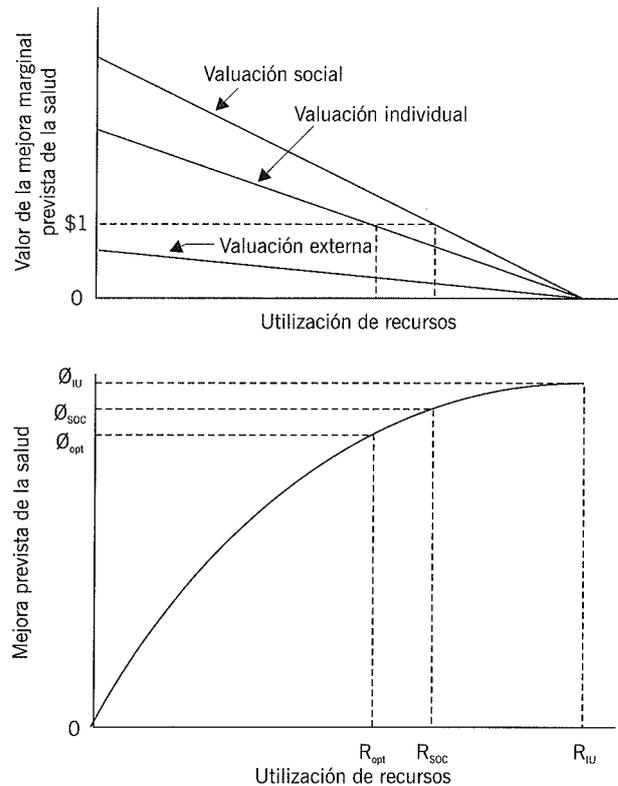
El interés de otras personas que no sean el paciente, entre ellas el médico, los familiares y amigos del paciente, y la sociedad en general, puede muy bien tenerse en cuenta al determinar cuál es la mejor estrategia de atención. La inclusión de los intereses de terceros en la salud del individuo conduce a otra posible definición de calidad: la "definición social"⁹. La definición social de calidad difiere de la definición individualizada porque incluye varias consideraciones, cada una de las cuales, a su vez, altera el resultado óptimo y el grado de utilización de recursos correspondiente.

En nuestro análisis de la definición individualizada de calidad nos basamos en la hipótesis de que cada persona paga todos los costos y recibe todos los beneficios de las decisiones relativas a las estrategias de atención. Una de las consecuencias de esta definición es que la especificación de calidad óptima varía según la capacidad del individuo para pagar los servicios de salud. Sin embargo, en muchos casos el individuo no paga todos los gastos ni recibe todos los beneficios.

El coste de la atención se comparte debido en parte a la existencia de programas de seguros para distribuir los riesgos y de programas de asistencia pública para corregir las desigualdades sociales. La presencia de estos mecanismos crea una divergencia entre el costo individual y el costo social y, en consecuencia, una disparidad entre lo que es óptimo para la sociedad y lo que es óptimo para el individuo. La disparidad surge porque, tal como han señalado varios investigadores, el enfermo tiene en cuenta únicamente el coste individual en la selección de la estrategia, mientras que la sociedad debe tener en cuenta el costo social total³⁷⁻³⁹. Por consiguiente, si la atención de la persona es financiada total o parcialmente por un seguro o una entidad de asistencia social, la persona tal vez prefiera una estrategia de atención más costosa y beneficiosa que la estrategia que la sociedad preferiría. En consecuencia, existe un conflicto de intereses, al menos a corto plazo, que coloca al médico en la difícil situación de tener que conciliar las preferencias inmediatas del individuo con los intereses mediatos de la sociedad⁹.

La presencia de beneficios compartidos o externos es un segundo factor que produce una diferencia entre lo que es óptimo para el individuo y lo que es óptimo para la sociedad. Cuan-

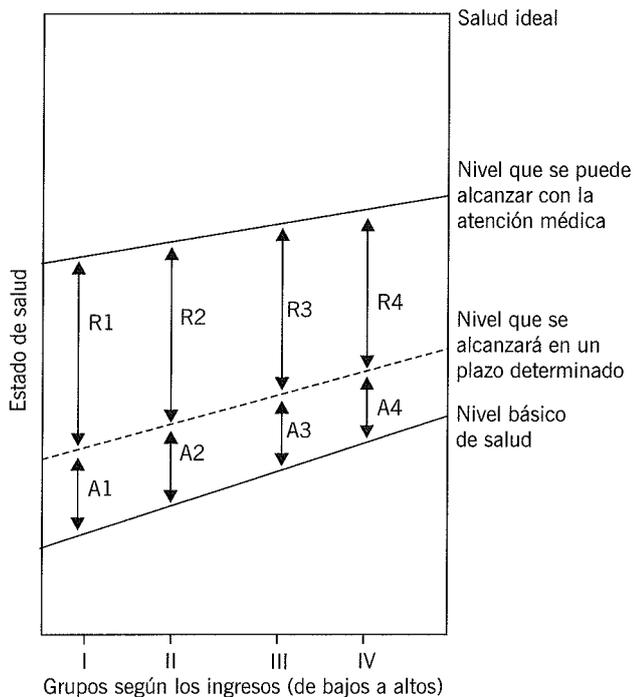
Figura 13. Efecto de la consideración de la valoración social en la definición de la estrategia óptima de atención



do un individuo recibe atención, además del beneficio que éste recibe, a menudo la sociedad también se beneficia. Por consiguiente, la valoración social de la atención consiste en la suma de la valoración individual y la valoración externa. El resultado es que la sociedad está dispuesta a pagar más por la atención que el individuo, y la magnitud de la disparidad depende de la magnitud de la valoración externa en comparación con la valoración individual.

La Figura 13 muestra la manera en que estos dos factores pueden influir en la utilización de recursos y en el resultado previsto en términos de salud. Ante la falta de beneficios externos y de un seguro, la manera óptima de proceder tanto para el individuo como para la sociedad se encuentra en R_{opt} . Ahora bien, si el individuo cuenta con un seguro y paga sólo una parte del coste de los recursos utilizados en el tratamiento, de manera tal que el coste individual es menor que el coste social, la utilización de recursos excederá de R_{opt} , que es la cantidad óptima desde el punto de vista social. En los casos limitados de suministro de atención médica gratuita a los pacientes, el gasto se acercará a R_{IU} . En esta situación, el coste social excede los beneficios sociales marginales. Por consiguiente, desde el punto de vista social, se destinan demasiados recursos al tratamiento, y la mejora adicional de la salud no justifica el gasto adicional.

Figura 14. Distribución hipotética del estado de salud en grupos de distintos ingresos



Por otra parte, si se tienen en cuenta los beneficios externos, la especificación de lo que es óptimo para la sociedad cambia de R_{opt} a R_{soc} . El problema radica en que, como ya se dijo, si el paciente no cuenta con un seguro o servicios de asistencia pública, preferirá gastar R_{opt} . Por lo tanto, desde el punto de vista social, se destinan muy pocos recursos al tratamiento del paciente, y la mejoría adicional resultante del estado de salud previsto justificaría la utilización de recursos adicionales.

Un tercer factor que produce una diferencia entre las especificaciones individualizada y social de la calidad es una preferencia social por una distribución determinada de la calidad o de incrementos del estado de salud entre distintos subgrupos de la población, distinguidos por uno o más factores de un conjunto que comprende la edad, el sexo, el origen étnico, la ubicación geográfica, los ingresos y otros factores similares. La Figura 14 ilustra alguna de las consideraciones pertinentes usando como ejemplo la distribución de la salud entre cuatro grupos de distintos ingresos.

Esta figura presenta varios postulados razonables. El primero es que el nivel actual de salud es inferior a lo que podría ser y está inversamente relacionado con los ingresos. El segundo postulado es que el suministro ilimitado de atención médica personal mejoraría el nivel general de salud, pero no lo llevaría hasta el nivel ideal, y reduciría las disparidades en el estado de salud relacionadas con los ingresos, pero no las eliminaría. La imposibilidad de que los servicios de salud personales alcancen

la salud ideal o eliminen totalmente las diferencias entre los grupos de distintos ingresos en un nivel de salud inferior al ideal surge de un tercer postulado, que reconoce la influencia de muchos otros factores en la salud y afirma que estos factores son sistemáticamente más adversos en los grupos de más bajos ingresos. No conocemos la importancia relativa a los servicios de salud personales en comparación con estas influencias adicionales en la salud. No queremos que nuestra figura efectúe una declaración en ese sentido ni nos atrevemos a definir lo que la salud ideal podría ser.

Sin embargo, postulamos que, de alguna manera, la sociedad evalúa el nivel actual de salud y su distribución en relación con un conjunto de valores y preferencias sociales. Después asigna recursos a fin de progresar en forma congruente con objetivos definidos socialmente. Por ejemplo, el objetivo podría ser alcanzar en el plazo de diez años un incremento equitativo de la salud de una magnitud específica en cada clase social definida por el nivel de ingresos. Por el contrario, bajo la influencia de un igualitarismo perfecto, la sociedad podría establecer el objetivo de disminuir en la práctica el nivel de salud de los grupos de mayores ingresos hasta que todos los grupos lleven una carga similar de mala salud. Suponemos que nuestra sociedad preferiría progresar en todos los grupos según los ingresos, mientras avanza más rápidamente en los grupos de menores ingresos a fin de que la distribución del estado de salud sea menos desigual.

Las repercusiones que esta decisión social tendrá en la asignación de recursos dependerá de la magnitud del cambio en el nivel y en la distribución de la salud que se desee alcanzar y en la facilidad con que el estado de salud pueda mejorar por medio de la atención personal. Es probable que la facilidad con que el estado de salud pueda mejorar no sea igual en todos los grupos según los ingresos, aunque no es fácil decir qué diferencias existen entre los grupos de distintos ingresos en este sentido. En cierta forma, las enfermedades de los grupos de ingresos más bajos tal vez sean más fáciles de evitar y tratar. En otros sentidos, tal vez sean más recalcitrantes, en parte debido a su índole y en parte porque, debido a la situación social de los pobres, atenderlos es más difícil y costoso. Cualquiera que sea la consecuencia neta de estos factores, el propósito es ofrecer la combinación y distribución de recursos que permita alcanzar de la manera más eficiente posible el cambio socialmente deseado en el nivel y la distribución actuales de la salud. En cierto sentido, se trata de una estrategia social de la atención que tiene en cuenta y abarca las estrategias de manejo clínico que usamos antes al formular nuestro modelo.

Suponemos que se puede determinar el coste de cualquier estrategia social de atención que se elija. Para cada grupo según los ingresos, el producto de la estrategia es la suma de todos los cambios en el estado de salud que puedan atribuirse a la atención personal en todos los miembros de ese grupo. Evidentemente, los costes incluyen todos los costes sociales, y los beneficios para la salud tienen en cuenta tanto la valuación individual como la externa. Lo que queda por especificar es un método para expresar el valor de los beneficios para la salud en los grupos de distintos ingresos.

Si la sociedad fuese neutral en cuanto a la distribución de la salud en relación con los ingresos, los beneficios para cada

grupo según los ingresos simplemente serían ponderados teniendo en cuenta el número de personas en cada grupo. Sin embargo, si existe una preferencia por un grado específico de redistribución entre los distintos grupos según los ingresos, tal como hemos postulado, la preferencia debe introducirse como ponderaciones que se atribuyen al resultado en términos de salud en cada grupo según los ingresos.

Donabedian ha ofrecido un ejemplo hipotético sencillo de la manera en que podría obtenerse un sistema de ponderaciones⁴⁰. Utilizando la notación de la Figura 14, la suma de los productos en términos de salud para todos los grupos de distintos ingresos sería:

$$A_1 P_1 (A_1 + R_1)^n + A_2 P_2 (A_2 + R_2)^n \\ + A_3 P_3 (A_3 + R_3)^n + A_4 P_4 (A_4 + R_4)^n$$

donde P es el número de personas de cada grupo y el exponente n es la función de la preferencia social. Si n= 0, la sociedad expresa neutralidad en cuanto a la distribución de los incrementos de la salud entre los grupos de distintos ingresos. Cuando n= 1, existe una preferencia inversamente relacionada con los ingresos, suponiendo que la distribución actual del estado de salud sea la indicada. Al aumentar el valor de n, aumenta rápidamente la pendiente de la preferencia a favor de la rectificación de la desigualdad actual de la distribución de la salud. Por supuesto, n podría ser también un número negativo, pero la mayoría de nosotros no querríamos vivir en una sociedad que acepte un valor n negativo, a pesar de la aprobación que aparentemente se da a este principio en la Sagrada Escritura⁴¹.

Comentarios finales

Empleando un concepto francamente hipotético, hemos llegado a una definición de calidad que establece una relación entre los incrementos del gasto total en atención personal de salud y los incrementos en un indicador del estado de salud que incluye una valoración personal y una valoración externa, y es ponderado aún más según la preferencia por una distribución social específica de la salud. El punto óptimo de la curva que describe esta relación corresponde a la estrategia social óptima de la atención.

Si la sociedad tiene en cuenta tanto los costes sociales como los costes individuales, tenderá a conformarse con una ganancia neta más pequeña en términos de salud que los individuos. Si tiene en cuenta tanto las valoraciones externas como las individuales, la sociedad tenderá a destinar más recursos a la atención que los individuos. Si el objetivo es redistribuir la salud entre los distintos subgrupos de la población, la sociedad quitará algo a algunos para dárselo a otros. En todos esos casos debemos recordar, una vez más, la índole contingente de la definición de calidad. Sin embargo, aunque la definición exacta de calidad es variable, consideramos que el marco conceptual y analítico básico es firme y coherente, y esperamos que también sea válido.

Nota

Nuestra selección de la representación gráfica del efecto de la atención médica en el estado de salud se benefició, cree-

mos, con la colaboración del profesor Donabedian y de T.M. Kashner, uno de sus alumnos doctorales, en la preparación de un modelo de utilización de servicios de salud que empleara un índice de salud basado directamente, en la medida de lo posible, en el trabajo de Rosser y Watts²⁹. Sin embargo, en este artículo hablamos solo en nuestro nombre y, por lo tanto, debemos absolver a los demás de cualquier complicidad en los errores que hayamos cometido.

Bibliografía

1. Cochrane AL. Effectiveness and efficiency: random reflections on health services. Londres: The Nuffield Provincial Hospitals Trust, 1972.
2. Ilich I. Medical nemesis: the expropriation of health. Nueva York: Pantheon Books, 1976.
3. Starr P. The politics of therapeutic nihilism. *Working Papers for a New Society* 1976;4:48.
4. Fuchs VR. Health care and the United States economic system - an essay in abnormal physiology. *Milbank Mem Fund Q* 1972;50: 211.
5. Fuchs VR. *Who shall live? Health, economics and social choice*. Nueva York: Basic Books, 1974.
6. Havinghurst CC, Blumstein JF. Coping with quality/cost trade-offs in medical care: the role of PSROs. *Northwestern University Law Review* 1975;70:6.
7. Donabedian A. Needed research in the assessment and monitoring of the quality of medical care. Hyattsville, Md.: National Center for Health Services Research, 1978. (DHEW publication No. [PHS] 78-3219).
8. Donabedian A. The quality of medical care: a concept in search of a definition. *J Fam Pract* 1979;9:277.
9. Donabedian A. *The definition of quality and approaches to its assessment*. Ann Arbor, Mich.: Health Administration Press, 1980.
10. Vuori H. Optimal and logical quality: two neglected aspects of the quality of health services. *Med Care* 1980;18:975.
11. Phelps CE. Benefit/cost analysis of quality assurance programs. In: Egdahl RH, Gertman PM, (eds). *Quality assurance in health care*. Germantown, Md.: Aspen Systems Corporation, 1976;289.
12. Brook RH, Davies-Avery A. Quality assurance and cost control in ambulatory care. In: Giebink GA, White NH, eds. *Ambulatory medical care quality assurance 1977: issues, directions, and applications*. La Jolla, Calif.: La Jolla Health Science Publications, 1977;46.
13. Brook RH, Davies-Avery A. Trade-off between cost and quality in ambulatory care. *Qual Rev Bull* 1977;3(11):4.
14. Fanshel S, Bush JW. A health-status index and its application to health services outcomes. *Operations Res* 1970;18:1021.
15. Breslow L. A quantitative approach to the World Health Organization definition of health: physical, mental and social well-being. *Int J Epidemiol* 1972;1:347.
16. Beloe NB, Breslow L, Hochstim JR. Measurement of physical health in a general population survey. *Am J Epidemiol* 1971; 93:328.

17. Berkman P. Measurement of mental health in a general population survey. *Am J Epidemiol* 1971;94:105.
18. Gilson BS, Gilson JS, Bergner M, et al. The sickness impact profile: development of an outcome measure of health care. *Am J Public Health* 1975;65:1304.
19. Bergner M, Bobbitt RA, Pollard WE, Martin DP, Gilson BS. The sickness impact profile: validation of a health status measure. *Med Care* 1976;14:57.
20. Pollard WE, Bobbitt RA, Bergner M, Martin DP, Gilson BS. The sickness impact profile: reliability of a health status measure. *Med Care* 1976;14:146.
21. Bergner M, Bobbitt RA, Carter WB, Gibson BS. The sickness impact profile: development and final revision of a health status measure. *Med Care* 1981;19:787.
22. Brook RH, Ware JE, Davies Avery A, et al. Overview of adult health status measures fielded in Rand's Health Insurance Study. *Med Care* 1979;17 (July Suppl):1.
23. Ware JE, Brook RH, Davies AR, Lohr KN. Choosing measures of health status for individuals in general populations. *Am J Public Health* 1981;71:620.
24. Stewart AL, Ware JE, Brook RH. Advances in the measurement of functional status: construction of aggregate indexes. *Med Care* 1981;19:473.
25. Carter WB, Bobbitt RA, Bergner M, Gilson BS. Validation of an interval scaling: the sickness impact profile. *Health Serv Res* 1976;11:516.
26. Patrick DL, Bush JW, Chen MM. Methods for measuring levels of well-being in a health status index. *Health Serv Res* 1973;8:228.
27. Blischke WR, Bush JW, Kaplan RM. Successive intervals analysis of preference measures in a health status index. *Health Serv Res* 1975;10:181.
28. Kaplan RM, Bush JW, Berry CC. Health status index: category rating versus magnitude estimation for measuring levels of well-being. *Med Care* 1979;17:501.
29. Rosser RM, Watts VC. The measurement of hospital output. *Int Epidemiol* 1972;1:361.
30. Torrance GW. Toward a utility theory foundation for health status index models. *Health Serv Res* 1976;11:349.
31. Blalock HM. *Social statistics*, second ed. New York: McGraw-Hill, 1972.
32. Sullivan DE. A single index of mortality and morbidity. *HSMHA Health Reports* 1971;86:347.
33. Brook RH, Davies-Avery A, Greenfield S, et al. Assessing the quality of medical care using outcome measures: an overview of the method. *Med Care* 1977;15(Sept Suppl):1.
34. Donabedian A. *Aspects of medical care administration: specifying requirements for health care*. Cambridge, Mass.: Harvard University Press, 1973.
35. Williamson JW. *Improving medical practice and health care: a bibliographic guide to information management in quality assurance and continuing education*. Cambridge, Mass.: Ballinger, 1977.
36. McNeil BJ, Weichselbaum R, Pauker SG. Fallacy of the five-year survival in lung cancer. *N Engl J Med* 1978;299:1397.
37. Arrow K. Uncertainty and the welfare economics of medical care. *Am Econ Rev* 1963;53:941.
38. Lees DS, Rice RG. Uncertainty and the welfare economics of medical care: comment. *Am Econ Rev* 1965;55:140.
39. Pauly MV. The economics of moral hazard: comment. *Am Econ Rev* 1968; 58:531.
40. Donabedian A. Models for organizing the delivery of personal health services and criteria for evaluating them. *Milbank Mem Fund Q* 1972; 50:103.
41. Evangelio según San Mateo 25:29.

La influencia de los factores socioeconómicos sobre la calidad asistencial

Donabedian A. The epidemiology of quality. *Inquiry* 1985;22:282-92

En el otoño de 1985 la revista *Inquiry* publicó un artículo del profesor Avedis Donabedian¹ titulado "La epidemiología de la Calidad". Este artículo, que el Profesor Donabedian había preparado por expreso encargo de los editores a partir de los datos que manejaba mientras escribía su libro "Métodos y Hallazgos en la Evaluación y Monitorización de la Calidad: un análisis ilustrado", recogía en forma sucinta, pero a la vez muy documentada, cómo factores sociales, organizativos, económicos y geográficos llegaban a afectar poderosamente a la calidad asistencial de la que son objeto los pacientes.

El Profesor Donabedian recurría, de este modo, a uno de los temas más clásicos y, probablemente al mismo tiempo, uno de los que constituyen por excelencia la razón de ser de los estudios sobre calidad: la variabilidad observada en los resultados y, en menor medida, en los procesos asistenciales. Su trabajo, plagado de referencias con las que documentaba su descripción panorámica y con la que llamaba la atención sobre los efectos de factores esencialmente no clínicos, constata un hecho hoy conocido pero, que en su momento, suponía una apuesta arriesgada y decidida.

Se ponía así de manifiesto cómo los pacientes con mayor renta era más improbable que padecieran comorbilidad tras someterse a cirugía (incluso menos probable el éxito como resultado de cirugía). Cómo era más fácil hallar casos de mortalidad evitable entre madres de recién nacidos con menores niveles de renta. Cómo la proporción de muertes perinatales evitables era mayor en hospitales municipales que en hospitales universitarios. Cómo el tipo de institución en la que se trabaja se relaciona directamente con el coste de los procedimientos diagnósticos. Cómo la presión asistencial reduce la cantidad de atenciones médicas que recibe el paciente y el cumplimiento de protocolos a igualdad de otros factores. O, sencillamente, cómo un mayor volumen de casos atendidos anualmente por el cirujano se asociaba a unos mejores resultados en cirugía.

El profesor Donabedian rehuyó en este trabajo entrar en el debate acerca de las razones o de los métodos apropiados para reducir esta variabilidad innecesaria y no deseable y recurrió al paralelismo con los estudios epidemiológicos para describir una serie de factores que encontraba afectan a los niveles de calidad asistencial de centros e instituciones sanitarias y que eran ajenos a la actividad clínica propiamente dicha.

Lugar de residencia, nivel de renta, procedencia social u otros factores de tipo organizacional de los centros sanitarios llegaban a condicionar y, sorprender al lector, con su influencia sobre la calidad asistencial; mientras que en otros casos (edad, procedencia rural, sexo u ocupación) no parecían ejercer ninguna influencia. Elogiar ahora la intuición del profesor Donabedian se antoja una tarea inútil pero con todo no podemos dejar de resaltar la oportuna reflexión que hace algo más de 16 años realizó y que estructuraba en seis grandes apartados.

Un primer gran apartado era casi obligado y hacía mención a la variabilidad no deseada fruto de la diferente localización de la práctica médica (por ejemplo variabilidad en resultados de cirugía en diferentes áreas geográficas de EE.UU., o número de histerectomías injustificadas en función de la ubicación geográfica del hospital).

Otro apartado lo dedicó al efecto que el entrenamiento, la experiencia, la especialización y la edad de los profesionales llegaba a tener sobre el resultado final. Años de práctica o volumen de casos atendidos (lo que incrementa notablemente la experiencia) se demostraban, junto con la especialización, predictores de niveles mayores de calidad.

Resulta también interesante la reflexión que el Prof. Donabedian realizó acerca de las condiciones organizativas de los centros y cómo estas condiciones afectan a la práctica médica. Dos temas de actualidad vigente, como el efecto de la presión asistencial y relación entre resultados y equipamientos para el diagnóstico y las terapias, ya fueron destacados y documentados por él.

Que la financiación y, concretamente los incentivos a los profesionales, llegan a afectar a la práctica y en qué dirección y que la disponibilidad de medios invita a un uso menos apropiado de los mismos, fueron otros de los tópicos abordados.

Una menor mortalidad en hospitales docentes (donde se forman los futuros profesionales) invitaba a la reflexión sobre la influencia de las características/perfil del centro sanitario en el nivel de calidad asistencial último que se oferta.

Por último, el Prof. Donabedian nos invita a considerar cómo distintos pacientes reciben diferente atención, tanto en cantidad como en cualidad, en función de determinadas características personales. No obstante, de todos los ejemplos que incluyó en su trabajo y tal y como señalaba el propio autor, lo menos aceptable es que para un mismo proceso y un mismo profesional se observen atenciones diferentes en función de esas características personales de los pacientes.

En palabras del Prof. Donabedian "igual acceso a un razonable nivel de atención a su vez equiparable, no necesariamente supone disfrutar de un nivel igual de salud". De ahí que la epidemiología de la calidad, pese al juego de palabras, sea un campo y una tarea que deba considerarse seriamente por todos políticos, gestores sanitarios, profesionales del sector y ciudadanos.

Bibliografía

1. Donabedian A. The epidemiology of quality. *Inquiry* 1985; 22:282-92.

José J. Mira
Universidad Miguel Hernández de Elche

Epidemiología de la calidad*

Avedis Donabedian

La epidemiología de la calidad es el estudio de la distribución de la calidad en un momento determinado y de los cambios en esa distribución a través del tiempo. Tiempo, lugar y persona, la tríada tradicional de la epidemiología descriptiva, es también aplicable al estudio de la calidad de la atención sanitaria. No obstante, la epidemiología de la calidad se diferencia en que posee, simultáneamente, dos grupos de población a los que pertenece: el de los proveedores sanitarios y el de los clientes. Esta dualidad, por lo menos, debería conferir a la epidemiología de la calidad unas características peculiares.

El conocimiento de la distribución de la calidad entre los profesionales de la salud (tanto los médicos individuales como los equipos, unidades de organización o instituciones enteras) es el producto de los estudios más comunes de calidad que arrojan luz a la relación entre la estructura, por una parte, y el proceso o resultado, por otra¹. A menudo, el propósito más inmediato de estos estudios es demostrar que ciertos atributos de los médicos se pueden utilizar para indicar una probabilidad mayor o menor de prestar una atención de calidad. De este modo, los atributos estructurales se convierten en medidas indirectas de calidad de la asistencia, así como indicadores de una posible asistencia deficiente y, por consiguiente, que merece una supervisión más frecuente. Más importante es aclarar la relación entre estructura y práctica asistencial, vital para la ciencia de la organización de asistencia sanitaria, mientras se pueda decir que esta ciencia existe. Debemos entender los determinantes de la práctica para, por ejemplo, diseñar un sistema de salud más eficaz.

La distribución de niveles específicos de calidad entre varios grupos de consumidores de asistencia es, por supuesto la medida definitiva de éxito o fracaso en la consecución de los objetivos sociales de un sistema de salud. Es necesario, por consiguiente, conocer a aquellos que reciben mejor o peor asistencia y las razones de esta desigualdad.

Vemos que la epidemiología de la calidad tiene funciones análogas a las de la epidemiología más conocida de los estados de salud: se trata, tanto de un método de investigación científico, como de una herramienta para la acción. Hasta el punto que los niveles de salud se utilizan en los estudios de calidad como medida de resultados, las dos epidemiologías, de hecho, se solapan, excepto en el hecho que para llevar a cabo la epidemiología de la calidad de las disparidades observadas en los niveles de salud deberían atribuirse a las diferencias en una atención previa.

El propósito de este estudio no es tratar la epidemiología de la calidad como método de investigación científica. Se trata simplemente de un estudio sobre lo que sabemos acerca de la relación entre los atributos de los profesionales sanitarios y de los clientes y la calidad de asistencia que prestan y reciben respectivamente. El lector puede consultar en otras fuentes sobre temas de definición y medición de calidad² y una evaluación detallada de los métodos utilizados en los estudios, de los que se seleccionó la mayor parte de la información en esta revisión³.

Incluso como vehículo para resultados empíricos, este estudio, aunque bastante detallado, es incompleto. En primer lugar, se han omitido los estudios que no ofrecen o no inducen a unos juicios relativamente directos sobre la idoneidad cuantitativa o cualitativa de la atención. Por consiguiente, la extensa bibliografía sobre modelos de utilización se ha excluido. La calidad se ha definido de un modo bastante restringido, por lo que la atención se centró en la asistencia técnica, facilitada principalmente por los médicos. La consiguiente omisión de la satisfacción del cliente es una exclusión especialmente limitativa. Finalmente, teniendo en cuenta el campo más limitado que hemos mencionado, la información recogida en este estudio pretende tener un carácter más ilustrativo que exhaustivo.

También es importante saber que no existen estudios de ámbito nacional sobre la calidad donde poder obtener una imagen nítida del nivel y de la distribución de la calidad, ni aquí ni en el extranjero. Para ese propósito, y también para interpretar las tendencias en calidad, nos basamos en los estudios de mortalidad, morbilidad, longevidad y sobre el uso de servicios que, en lo que respecta a sus implicaciones en la calidad de la asistencia, son extremadamente difíciles de interpretar. Contamos con un conjunto bastante amplio de estudios, utilizando métodos que varían en precisión y rigor, para evaluar diversos aspectos de la calidad de la atención, facilitada por médicos seleccionados en colectivos circunscritos en periodos distintos. Como consecuencia de esto, poco puede decirse de los niveles de calidad en el conjunto de los Estados Unidos, además, vemos que invariablemente, la práctica asistencial parece estar muy alejada de los criterios y estándares utilizados para juzgar la calidad de la asistencia. Como a medida que vaya avanzando la ciencia médica, estos criterios y estándares van a ser cada vez más rigurosos y habrá poco que aprender de tendencias seculares a partir de comparar los estudios actuales con los del pasado, aunque en otros campos, la comparación parecería más

*Donabedian A. The epidemiology of quality. *Inquiry* 1985;22:282-92

acertada. Por estas mismas razones, lo que se ha dicho acerca de los factores que influyen en la calidad también debería ser tratado con cautela, aunque en este caso, estamos más seguros de la aplicabilidad general de las asociaciones observadas, especialmente, cuando aparecen repetidamente y parecen coherentes con lo que conocemos del sistema de salud en general.

Variabilidad y localización

La distribución de las conductas que, directa o indirectamente, denotan calidad, revela un sorprendente grado de variabilidad entre las áreas geográficas, los médicos y las instituciones. En los últimos años, se ha prestado atención especial a las diferencias poco entendidas en la incidencia de los procedimientos quirúrgicos entre pequeñas áreas de los Estados Unidos⁴, dentro de otros países⁵ y entre países⁶. También existen diferencias considerables entre médicos e instituciones, en cualquier área geográfica delimitada, e incluso entre médicos de dentro de una institución. Se ha observado que el cumplimiento de los criterios implícitos o explícitos de una buena asistencia varía entre médicos generales⁷ e internistas⁸ que ejercen en un área. Lo mismo se puede aplicar a los médicos dentro de un conjunto de prácticas asociadas en grupo⁹. Existen diferencias notables en el coste de la asistencia prestada por el médico en las prácticas en grupo¹⁰ y por parte de médicos y otros profesionales en hospitales universitarios¹¹. El número y el coste de los procedimientos de diagnóstico varían considerablemente entre los médicos que ejercen en un centro¹², un hospital¹³ o una práctica en grupo¹⁴. La tendencia a la extirpación de tejidos normales en apendicetomías, varía sorprendentemente entre los médicos de un hospital y también entre el personal de hospitales distintos¹⁵.

Los hospitales y las instituciones de atención ambulatoria también muestran diferencias respecto a otros aspectos de la calidad. Por ejemplo, los hospitales difieren enormemente en el porcentaje de histerectomías injustificadas practicadas¹⁶, en la incidencia de muertes y complicaciones postoperatorias¹⁷, en la prevención de mortalidad¹⁸ materna y en el cumplimiento de criterios implícitos o explícitos de su personal respecto a la utilización y calidad adecuada¹⁹.

La calidad de asistencia que facilitan los centros de asistencia ambulatoria organizados varía de manera similar²⁰.

Por grandes que puedan resultar las diferencias en la práctica médica, la variabilidad no es casual. Las prácticas cuestionables, por el contrario, están muy concentradas o localizadas. Ha podido observarse, por ejemplo, que un porcentaje relativamente pequeño de médicos prescribe un porcentaje bastante elevado de medicamentos²¹ e inyectables²² discutibles. La concentración de tonsilectomías²³ y de otros procedimientos quirúrgicos²⁴ en pocos cirujanos, levanta sospechas de que los responsables de la mayor parte de las intervenciones, se pueden estar excediendo, mientras que los que realizan demasiado pocas, son incapaces de mantener sus habilidades quirúrgicas.

La variabilidad y la localización son las necesidades gemelas de la epidemiología. Lo que queda es identificar los atributos con los que se asocian el uso adecuado y la calidad.

Formación, experiencia, especialización y edad de los médicos

Los atributos principales que influyen en la actuación médica (más allá de la posible influencia de rasgos de personalidad, un tema del que no poseo información) parecen ser la formación, la experiencia y la especialización.

El efecto de la formación se ha ilustrado repetidamente en la bibliografía. La práctica general suele ser de mucha mayor calidad entre aquellos médicos que en su época habían sido mejores estudiantes²⁵, que fueron educados en distintas clases de facultades de medicina²⁶ y que se sometieron a ciertos tipos de formación²⁷ durante períodos más prolongados. Este efecto se ve reforzado cuando coexisten dos o más características de este tipo²⁸. No obstante, los médicos de medicina general, como grupo, se observa que mantienen unos estándares de práctica médica más bajos, en comparación con sus colegas más especializados, como mínimo cuando se les juzga en función de criterios que, normalmente, se ajustan más a las opiniones de los especialistas.

Los especialistas suelen realizar mejor su práctica profesional que los médicos de medicina general en muchos aspectos de la atención ambulatoria y hospitalaria. Los especialistas tienen un mayor criterio al realizar pruebas de rayos x²⁹, realizan menos inyectables que pueden ser cuestionables³⁰, utilizan el hospital de manera más adecuada³¹, cumplen más a menudo con los criterios implícitos o explícitos de la asistencia hospitalaria³² y ambulatoria³³, y son responsables de un menor número de muertes perinatales "evitables"³⁴. Existen pruebas, además, de que el efecto de la especialización está relacionado con los campos de actuación, aquellos casos que atañen estrictamente a su área de especialización³⁵. Además, cuanto más especializada es una organización, mejor desempeña su labor de asistencia³⁶.

La formación y la especialización están, por supuesto, estrechamente relacionadas con la experiencia. Existen dos componentes adicionales de la experiencia: 1. años de práctica profesional y 2. número de casos similares tratados, o intervenciones similares realizadas cada año. Respecto al segundo de ellos (el volumen de la asistencia), existen indicios de que los mayores hospitales prestan una mejor asistencia técnica³⁷, aunque no necesariamente, una asistencia que satisfaga al paciente³⁸. No obstante, el tamaño de la institución es una característica con distintas connotaciones; la experiencia en un mayor número de casos similares, tanto a través de médicos individuales, como a través del conjunto del personal del hospital. Es interesante descubrir, por tanto, que la mortalidad derivada de ciertas intervenciones quirúrgicas³⁹, es inferior en los hospitales donde se realiza un mayor número de estos procedimientos, y que las muertes maternas evitables suelen reducirse en los hospitales que realizan un mayor número de partos⁴⁰. Contrariamente, quizás atribuible a factores adversos asociados al tamaño del hospital, existe una asociación positiva entre el número anual de nacimientos y la mortalidad maternal y neonatal, según una muestra realizada en hospitales nacionales⁴¹.

La influencia de la experiencia, que parece aumentar con una práctica más prolongada, tiene contrapartidas, ya que con el paso del tiempo, el conocimiento médico se va quedando más obsoleto.

Consecuencia de ello, la edad avanzada parece ir siempre asociada a unos niveles más bajos de práctica asistencial. Aunque no está claro si esto es bueno o no, los médicos de mayor edad emplean menos pruebas de laboratorio y rayos x⁴². Otro tipo de observaciones, más indicativas de la calidad, muestran que los médicos de más edad y con más años de experiencia utilizan las pruebas de rayos x con menos rigor⁴³; emplean más a menudo inyectables⁴⁴ que pueden ser cuestionables y su trabajo como médicos de medicina general en la comunidad⁴⁵, como médicos de familia en la práctica en grupo⁴⁶ y como médicos en diversos centros de asistencia ambulatoria⁴⁷ es peor; y prestan una asistencia hospitalaria que probablemente se ajusta menos a unos criterios explícitos prestablecidos que la asistencia facilitada por médicos con menos años de experiencia⁴⁸. Un estudio mostraba un modelo que contradice esto último, los mejores pediatras de la comunidad resultaron ser aquellos que habían estado ejerciendo o menos de 12 años o más de 40; aunque en general, cuanto más tiempo habían ejercido los médicos, menor resultaba su adhesión a los criterios de atención ambulatoria⁴⁹.

Puesto que estos estudios sobre la relación entre edad y práctica asistencial cuentan con un diseño transversal, habiendo observado únicamente diferencias contemporáneas entre los médicos, no podemos saber hasta qué punto la diferencia entre los jóvenes y los mayores se debe a un fallo de estos últimos, en la adquisición de nuevos conocimientos o aptitudes y en qué medida, se debe a la pérdida de conocimientos o aptitudes originales. Es probable también, que las comparaciones entre pediatras y médicos de medicina general impliquen que en determinadas circunstancias, que ahora sólo son conjeturas, el equilibrio de estas dos tendencias (la adquisición y pérdida de competencia) permanezca favorable, al menos hasta que una edad demasiado elevada se cobre su peaje.

Condiciones de la práctica en consultorio

La actuación de los médicos también está influenciada por las condiciones de su propia práctica y por las organizaciones o instituciones en que trabajan. Por ejemplo, los médicos de medicina general en la zona rural de Carolina del Norte de principios de los años 50, realizaban mejor su trabajo cuando podían contar con una gran variedad de equipamiento médico (microscopio, centrifugadora clínica, electrocardiograma, un aparato para determinar las tasas metabólicas basales y un calorímetro fotoeléctrico) en sus consultorios⁵⁰. La atención a los pacientes en consultas mediante cita previa comportaba unos mejores niveles de actuación por parte de los médicos de medicina general en las zonas rurales de Carolina del Norte⁵¹, los médicos de medicina general en las provincias canadienses de Ontario y Nueva Escocia⁵² y por parte de médicos de familia en una asociación de práctica en grupo en la ciudad de Nueva York y alrededores⁵³. Los médicos de familia de este último sitio prestaban mejor asistencia si contaban con un ayudante administrativo a tiempo total o habían utilizado una historia clínica de mayor formato⁵⁴.

Los médicos demasiado ocupados pueden que tiendan a prestar una asistencia menos completa (según las historias estudiadas) cuando el número atendido de pacientes por hora se

eleva demasiado⁵⁵. Cuanto mayor es el volumen de trabajo, mayores indicios de menor cumplimiento de los criterios para la elaboración de historias clínicas y para el examen físico, aunque sí existe un mayor cumplimiento de los criterios para la petición de pruebas de laboratorio⁵⁶. Esto sugiere que esto último podría sustituir a lo anterior.

Financiación y organización de la asistencia

Existen también indicios, no siempre del todo convincentes, de que los incentivos económicos pueden influir en la práctica. Los médicos que cuentan con su propio equipo de rayos x tienen más tendencia a utilizarlo y a hacerlo de forma menos apropiada⁵⁷. Se realizan más intervenciones quirúrgicas cuando los médicos no pueden o no desean cargar más costes de los que marca la tarifa de la aseguradora⁵⁸, las apendicetomías son probablemente menos justificadas cuando la estancia en el hospital está financiada por Plan de Blue Cross⁵⁹, los pacientes que han colaborado en el pago de su estancia hospitalaria experimentan unas estancias más largas y más frecuentes y con menos frecuencia tienen estancias cortas que los pacientes que deben costearse ellos la factura⁶⁰. Estas observaciones se hacen más evidentes con las evidencias que muestran que cuando el pre-pago se combina con las prácticas en grupo, se produce una utilización de servicios más parsimoniosa y quizás más adecuada⁶¹.

La participación en un grupo, clínica o institución similar podría alterar radicalmente los términos y las condiciones de la práctica, aportando los cambios correspondientes en la calidad de la asistencia técnica. Se ha podido observar, por ejemplo, que la calidad de la atención ambulatoria para niños es mejor en las consultas externas del hospital que en la práctica de consultorio, al menos si lo juzgamos a partir de los registros de la historia clínica⁶². En la zona rural de Carolina del Norte, los médicos de medicina general que ejercían en asociaciones de dos o tres miembros, prestaban una mejor asistencia, tal vez porque estaban más cualificados que la media en cuanto a estudios y formación, eran más jóvenes y se habían organizado y equipado para favorecer una buena labor asistencial⁶³. No se observaron relaciones de este tipo entre calidad y trabajo en asociaciones de dos o tres miembros en médicos de medicina general en las provincias canadienses de Ontario y Nueva Escocia⁶⁴.

En este sentido, un estudio más reciente en Hawaii sugiere que los médicos en grupos de especialidades múltiples no demostraron facilitar una atención ambulatoria mejor que la prestada por médicos individuales⁶⁵. Los médicos de esos grupos, sin embargo, mostraron claramente una mejor utilización del hospital y prestaban mejor asistencia hospitalaria que los médicos que actuaban individualmente⁶⁶. No obstante, la relación entre pertenencia a grupo y la naturaleza de la atención hospitalaria se explicaba por el hecho de que los grupos empleaban a especialistas que limitaban su práctica a sus especialidades respectivas⁶⁷.

Sería interesante e importante saber si la práctica en grupos produce un efecto en la calidad a través de mecanismos diferentes de la mera selección de médicos que participan en el grupo, aunque tenga su importancia. Otros ejemplos de formas

más sutiles de influencia se pueden extraer a partir de la observación de que la calidad de la atención prestada por médicos de familia en un conjunto de prácticas asociadas en grupo era mejor si los médicos llevaban más tiempo afiliados al grupo, concedían más tiempo al grupo y se pasaban más tiempo en las instalaciones⁶⁸. La importancia de una afiliación activa al grupo se confirma, si observamos que la calidad de la atención ambulatoria resultaba mejor cuando los médicos tenían este compromiso, cuando este compromiso tenía un mayor rango y cuando se realizaba un mayor número de visitas hospitalarias⁶⁹. No obstante, para destacar la complejidad de la relación hospital y médico, no existió relación alguna, en la zona rural de Carolina del Norte, entre la calidad de la asistencia de consultorio y el grado de utilización del hospital por parte del médico, aunque los mejores doctores se mostraban más selectivos y aptos en la utilización de los servicios hospitalarios⁷⁰.

Existen algunas evidencias sobre la relación entre el acceso de los pacientes a una práctica organizada y el resultado de la asistencia. La incidencia de primeros episodios de fiebre reumática se ha reducido después de haberse fundado un centro asistencial completo en una zona de nivel socioeconómico bajo⁷¹. Los afiliados a un plan de asistencia de grupo con pago anticipado recibieron una mayor asistencia prenatal y mostraron una tasa inferior de mortalidad neonatal que la ofrecida en una muestra de población general, siendo mayores las ventajas para la población de color⁷². Aunque un estudio más reciente en otra comunidad, muestra que los afiliados a una asistencia de grupo recibían menor asistencia prenatal, aunque mantenían el mismo porcentaje de mortalidad neonatal, que hizo que los investigadores sugirieran un mayor grado de eficiencia en la consecución de unos resultados determinados⁷³.

Características del hospital

De todas las organizaciones que prestan asistencia sanitaria, los hospitales son quizás los que mejor pueden controlar la calidad de la atención que realizan. Las características de los hospitales, individuales o asociadas a las características de los médicos que trabajan en ellos, podrían tener una gran influencia sobre la calidad de la asistencia. Ya he aludido a los posibles efectos del tamaño del hospital. Mucho más importante parece ser la función educativa del hospital, de manera especial cuando esa función va acompañada de una estrecha afiliación a una facultad de medicina.

Los hospitales universitarios suelen utilizarse más adecuadamente⁷⁴. Las apendicetomías suelen tener más justificación en este tipo de hospitales⁷⁵. La asistencia prestada en los hospitales de formación cumple más a menudo con los criterios implícitos y explícitos de calidad, al menos según muestra la historia clínica⁷⁶. Las tasas de defunciones resultaron más bajas en los hospitales universitarios de Inglaterra y Gales⁷⁷ y las muertes perinatales evitables fueron inferiores en los hospitales universitarios de la ciudad de Nueva York, en comparación con sus referentes voluntarios o municipales, que no tenían funciones formativas⁷⁸.

Existen algunas excepciones inquietantes dentro de esta armonía general de opinión. En un estudio nacional de asistencia de obstétrica, las tasas de mortalidad maternal, morta-

lidad neonatal y el porcentaje de abortos fue mucho mayor en hospitales universitarios, por otro lado comprensible ya que no se ajustaron las diferencias en "case mix"⁷⁹. Esta misma explicación no sirve para otros estudios que, habiendo corregido ampliamente las diferencias entre los pacientes, siguieron mostrando que los efectos adversos postoperatorios y el ser hospital universitario (esto último medido por la vinculación con la universidad por la presencia de un programa de formación de residentes o por la razón de número de personal con respecto al número de camas) no estaban relacionados⁸⁰ o estaban relacionados positivamente⁸¹.

La información sobre la calidad de la atención en los hospitales con ánimo de lucro es fragmentaria; además, es difícil de interpretar ya que estos hospitales difieren mucho entre ellos y están muy influenciados por las circunstancias de tiempo y lugar.

Tras ajustar las diferencias en el tiempo de estancia, el porcentaje de defunciones de pacientes internados en hospitales con ánimo de lucro resultó más elevado que el de los pacientes en hospitales no lucrativos en Los Angeles County en 1964⁸², aunque no fue claramente así en la ciudad de Nueva York en 1971⁸³. Contrariamente, durante 1950-1951, se produjeron menos muertes perinatales evitables en los hospitales de propiedad de la ciudad de Nueva York, que en cualquier otro tipo de institución, exceptuando los hospitales no lucrativos.

Aunque una vez más, en la ciudad de Nueva York, la asistencia en hospitales con ánimo de lucro⁸⁴, al menos para un cierto segmento de población durante 1961 o 1962, resultó mucho más inadecuada y de una calidad "menos que óptima", especialmente en manos de médicos a quienes sólo los hospitales con ánimo de lucro podrían conceder privilegios⁸⁵.

Vista la gran importancia que se le da a la acreditación de hospitales, es de destacar lo poco que se conoce acerca de su relación con la calidad de la asistencia. Esto puede ser debido a que todos los mejores hospitales ya están acreditados, por lo que la contribución que podría atribuirse específicamente a la acreditación es difícil de discernir. En un estudio de ámbito nacional, no se apreció una relación significativa entre la situación de acreditación y la mortalidad postoperatoria de tres procedimientos después que las diferencias de case mix habían sido corregidas⁸⁶. Entre otras informaciones sobre el tema, la mortalidad hospitalaria en función del tiempo de estancia es inferior en hospitales acreditados⁸⁷ y la acreditación parece estar relacionada con una mayor calidad en comparación con hospitales sin ánimo de lucro similares, aunque no lo está si la comparación se establece entre hospitales privados similares acreditados y no acreditados⁸⁸. Un estudio realizado en 30 hospitales del área de Chicago, todos acreditados, reveló que las valoraciones mejor puntuadas de los departamentos médicos por parte de los evaluadores de la *Joint Commission on Accreditation of Hospitals* estaban relacionadas con unos porcentajes de fallecimientos inferiores, después de ajustar las diferencias en el tiempo de estancia y con una mayor reputación para los hospitales entre las personas con un buen nivel de información⁸⁹.

No hay duda de que los hospitales influyen sobre la calidad, en parte a través de los servicios y del equipo médico que facilitan y, en parte mediante la selección de los médicos y otro

personal empleado. No obstante, parece que existe otro efecto a través de controles organizativos. En cualquier caso, existe una interacción entre las características de los hospitales y los atributos de los médicos que ejercen en ellos. La cualificación formal de los médicos aparece como un indicador de calidad de menor importancia en los hospitales universitarios que en los otros. Contrariamente, cuanto menos capaz sea el hospital y su personal médico de controlar la práctica de la medicina en el hospital, más probable será que la atención que reciba el paciente dependa de las cualificaciones de los médicos individuales⁹⁰. Independientemente de la afiliación universitaria, la calidad de la atención técnica, parece ser mejor si está en manos de médicos con una gran preparación, en hospitales altamente organizados y peor si está en manos de médicos menos preparados con formas de organización menos desarrolladas⁹¹.

Un gran número de estudios indagan en la teoría de las organizaciones para explorar la relación entre la calidad y los detalles de la estructura y el comportamiento de la organización. Estos estudios ofrecen posibles explicaciones para algunas de las correlaciones más evidentes que hemos conocido, aunque revelan también nuevas asociaciones que no habíamos sospechado y que, a veces, vemos difíciles de entender o creer⁹². Creemos que, por ejemplo, aunque el tamaño del hospital, la situación formativa o un mayor porcentaje de personal cualificado no guarde relación muchas veces con unos mejores resultados (tanto quirúrgicos como médicos), estos resultados parecen mejores cuando existen niveles más altos de coordinación en el hospital en conjunto, en las unidades de asistencia al paciente o en los quirófanos⁹³ y cuando existe un mayor control sobre los privilegios y las actividades del personal quirúrgico⁹⁴. Por otro lado, un estudio reveló inesperadamente que cuando el personal de enfermería contaba con un mayor grado de poder en su propio "dominio", la mortalidad en lugar de mejorar, empeoraba⁹⁵. Aunque dudamos de este descubrimiento, a menos que signifique una menor comunicación entre médicos y enfermeros, creemos que se gana una nueva perspectiva sobre la relación entre edad y actuación, cuando se nos dice que los hospitales que emplean un mayor número de médicos con mucho tiempo ejerciendo, se trata también de hospitales que se caracterizan por una mayor relajación en la coordinación y el control⁹⁶.

Características de los clientes

Hasta ahora, nos hemos ocupado de la distribución de la calidad técnica entre los proveedores. Se sabe mucho menos acerca de su distribución entre los clientes, aunque debería de tratarse de un tema de máxima importancia para la política pública⁹⁷.

Las diferencias de calidad de la atención prestada a diversas categorías de personas se pueden dar de una o dos maneras. Diferente gente puede tener distintos grados de acceso a diversos tipos de médicos y de instituciones. Esta es una manera de diferenciación social que parecemos dispuestos a tolerar, aun cuando lo consideramos desafortunado. No somos tan capaces de contemplar la segunda posibilidad: que los pacientes con necesidades de asistencia técnica similares sean tratados de manera distinta cuando los atiende el mismo médico o la misma institución, a menos que la diferencia radique única-

mente en temas de confort. Esta reticencia puede explicar en parte, nuestro poco interés o poca voluntad o incapacidad para estudiar el tema, con el resultado de que se conoce muy poco sobre ello. Se sabe más acerca de un tercer factor que influye sobre la calidad de la asistencia, supuestamente, la capacidad o voluntad de los clientes de participar para influir en las recomendaciones del médico y en llevarlas a cabo. Esta revisión, no obstante, no incluye los estudios sobre la participación y cumplimiento del paciente.

La variabilidad en la cantidad y calidad de la asistencia ya mencionada entre los proveedores debe tener su equivalente entre las personas que utilizan los servicios. Esperamos ver, por consiguiente, las diferencias correspondientes entre los que residen en distintas áreas geográficas o aquellos que son clientes de diferentes hospitales y médicos. Los clientes de cualquier grupo de médicos pueden llegar a recibir una calidad de asistencia muy distinta⁹⁸.

En cuanto a la relación entre los atributos del cliente y la calidad de la atención, existen datos aislados sobre los efectos del lugar de residencia, edad, sexo, raza, ocupación, estar en posesión de seguro y los índices de la situación económica.

Cabe esperar que los que residen en zonas rurales experimenten las consecuencias adversas de un acceso limitado a la asistencia, especialmente a la asistencia de médicos especializados y a la asistencia en grandes hospitales universitarios. Por ejemplo, el uso indebido de antibióticos suele darse más en las áreas rurales, aun teniendo en cuenta la situación de especialización del médico⁹⁹. Los procedimientos quirúrgicos no se realizan tan a menudo con la población residente en áreas menos urbanizadas¹⁰⁰. Aunque considerando la posibilidad de intervenciones quirúrgicas no necesarias, el estar menos sujeto a intervenciones no tiene por qué ser algo malo. Al menos un estudio ha constatado una relación entre la residencia urbana y un número más elevado de apendicetomías y una mayor tasa de mortalidad derivada de la apendicitis¹⁰¹.

Las pruebas referentes al posible efecto de la edad y del sexo son fragmentarias y poco concluyentes. Por una parte, no parece existir una relación consistente entre la edad de los pacientes y la calidad, tanto de la atención hospitalaria como de la atención ambulatoria en Hawaii¹⁰². Por otra parte, la calidad de los servicios de enfermería, en un grupo seleccionado de hospitales urbanos, ha resultado mejor en los jóvenes que en las jóvenes, y mejor en las mujeres mayores que en los hombres mayores¹⁰³. En un estudio internacional, se demostró que las mujeres eran más propensas a la apendicetomía y que la intervención mostraba un tejido normal con mayor frecuencia, aunque esto tendría explicación en el hecho de que existen mayores dificultades en diagnosticar apendicitis en las mujeres¹⁰⁴.

Las comparaciones basadas en diferencias raciales son igualmente escasas y poco concluyentes. No se apreció relación entre origen étnico y la calidad de la asistencia hospitalaria en Hawaii, ni tampoco en la atención hospitalaria, al menos en cuanto al cumplimiento de estándares técnicos, influenciados en gran manera si el paciente recibía atenciones de un médico de origen étnico similar o distinto¹⁰⁵. Hawaii podría ser un caso atípico en este sentido. Otros estudios han descubierto que las intervenciones quirúrgicas son menos frecuentes entre la población que no es blanca¹⁰⁶, que la población no blanca puede estar recibiendo una atención de enfermería de peor calidad¹⁰⁷.

y que en caso de intervenciones, las realizan más a menudo personal residente que cirujanos con más experiencia o más cualificados¹⁰⁸. No puede concluirse, no obstante, que la asistencia que presta el personal residente sea necesariamente inferior. Un estudio mostró que el personal residente realizaba muy bien su labor¹⁰⁹, mientras que en otro estudio su actuación fue catalogada de bastante pobre¹¹⁰.

Las diferencias ocupacionales son interesantes, en parte porque la ocupación es un indicador importante de la situación socioeconómica (como lo es, desgraciadamente, el origen étnico) y, en parte, porque la utilización de servicios por parte de los médicos y familiares puede usarse como estándar de atención¹¹¹. Se ha observado, por ejemplo, que las esposas de médicos corren más riesgo de sufrir histerectomías que el resto de mujeres en general¹¹². Esto puede significar también, sin embargo, que las mujeres de los médicos están más expuestas a intervenciones quirúrgicas innecesarias. Se sabe que las histerectomías se suelen proponer y realizar frecuentemente sin estar del todo justificadas¹¹³. Un estudio realizado en la ciudad alemana de Hannover muestra que los trabajadores no manuales o administrativos se someten a más apendicetomías, aunque el tejido que se extrae muestra, a menudo, inflamaciones menos graves¹¹⁴. Otro aspecto sería la observación de que las personas relacionadas con la medicina o la enfermería pueden experimentar más riesgos dada su mayor predisposición para someterse a apendicetomías que, según se demuestra, no están justificadas¹¹⁵.

Estas observaciones sugieren que un mayor acceso a la asistencia puede tener consecuencias positivas y negativas. Este principio puede aplicarse cuando la cobertura de un seguro médico privado o de un programa gubernamental mejora la capacidad de pago por la asistencia. Se ha observado que esta cobertura reduce la infrutilización del hospital y, al mismo tiempo, incrementa su sobreutilización¹¹⁶. Las apendicetomías para pacientes cuyas estancias hospitalarias estaban financiadas por Blue Cross Plans mostraban un tejido normal o cuestionable en un mayor porcentaje de los casos¹¹⁷.

En algunos estudios, los resultados de asistencia se han utilizado para mostrarnos algunos aspectos de la influencia de los factores sociales y económicos sobre la calidad de la atención y sus consecuencias. Se ha dicho, por ejemplo, que los pacientes con mayores ingresos tienen menos probabilidades de sufrir complicaciones o muerte postoperatoria, incluso después de ajustar por otros factores¹¹⁸. Las madres pertenecientes a unas categorías socioeconómicas inferiores en función de la renta media, han mostrado unos mayores porcentajes de mortalidad materna evitable, diferencia ampliamente atribuible a que la madre no puede obtener una asistencia adecuada¹¹⁹. La proporción de muertes perinatales consideradas evitables es mayor cuando el parto tiene lugar en un hospital municipal que en un hospital no lucrativo, o cuando la madre y el niño reciben cuidados en una sala de hospital y no en un servicio privado de hospital¹²⁰. Cuando se mejora la organización de la asistencia, puede darse una mejora asociada en la asistencia prenatal y una reducción de la mortalidad neonatal, mejoras relativas que son mayores en el caso de los más desfavorecidos. Aunque esta desventaja no se remedia del todo¹²¹. Ni siquiera un acceso prácticamente idéntico a una asistencia casi igual deriva necesariamente en un mismo nivel de salud.

Eso no significa, por supuesto, que debamos de perder la esperanza; sólo que debemos intentar mejorar en lo posible y en un ámbito más amplio.

Bibliografía y notas*

Este artículo está buscado en el epílogo del libro del autor *The Methods and Findings of Quality Assessment and Monitoring: An Illustrated Analysis*, published by the Health Administration Press. Copyright to the book is held by the Regents of the University of Michigan. The work was supported by the Commonwealth Fund, the Carnegie Corporation of New York, the Milbank Memorial Fund, the W. K. Kellogg Foundation, and the National Center for Health Services Research. Los puntos de vista del autor no tienen porque coincidir con el de sus sponsors.

1. Donabedian A. *Explorations in Quality Assessment and Monitoring*, vol. 1: *The Definition of Quality and Approaches to Its Assessment*. Ann Arbor, MI: Health Administration Press, 1980.
2. Ibid, Donabedian A. *Explorations in Quality Assessment and Monitoring*, vol. 2: *The Criteria and Standards of Quality*. Ann Arbor, MI: Health Administration Press, 1982.
3. Donabedian A. *Explorations in Quality Assessment and Monitoring*, vol. 3: *The Methods and Findings of Quality Assessment and Monitoring: An Illustrated Analysis*. Ann Arbor, MI: Health Administration Press, 1985.
4. Lembcke PA. "Measuring the Quality of Medical Care Through Vital Statistics Based on Hospital Service Areas, 1: Comparative Study of Appendectomy Rates," *American Journal of Public Health* 1952;42:276-86; Lewis CE. "Variations in the Incidence of Surgery". *New England Journal of Medicine* 1969;281: 880-4; Wennberg J, Gittelsohn A. "Small Area Variations in Health Care Delivery". *Science* 1973;182:1102-8; Gittelsohn AM, Wennberg JE. "On the Incidence of Tonsillectomy and Other Common Surgical Procedures," in *Costs, Risks, and Benefits of Surgery*, ed. Bunker JP, Barnes BA, Mosteller F. New York: Oxford University Press, 1977; Griffith JR, Restuccia JD, Tedeschi PJ, Wilson PA, Zuckerman HS. Measuring Community Hospital Services in Michigan. *Health Services Research* 1981;16:135-60.
5. Stockwell H, Vayda E. Variations in Surgery in Ontario. *Medical Care* 1979;17:390-6.
6. Bunker JP. Surgical Manpower: A Comparison of Operations and Surgeons in the United States and in England and Wales. *New England Journal of Medicine* 1970;282:135-44; Vayda E. A Comparison of Surgical Rates in Canada and in England and Wales. *New England Journal of Medicine* 1973;289:1224-9; McPherson K, Wennberg JE, Hovind OB, Clifford P. Small-Area Variations in the Use of Common Surgical Procedures: An International Comparison of New England, England, and Norway. *New England Journal of Medicine* 1982;307:1310-4.
7. Peterson OL, Andrews LP, Spain RS, Greenberg BG. An Analytical Study of North Carolina General Practice, 1953-1954. *Journal of Medical Education* (pt. 2) 1956;31:1-165; Clute KF. *The General Practitioner: A Study of Medical Education and Practice in Ontario and Nova Scotia*. Toronto: University of Toronto Press, 1963.

*Las referencias bibliográficas se reproducen en su formato original con las notas del autor

8. Hulka BS, Romm FJ, Parkerson GR, Jr, Russell IT, Clapp NE, Johnson FS. Peer Review in Ambulatory Care: Use of Explicit Criteria and Implicit Judgments. *Medical Care* 1979;17:1-73.
9. Morehead MA. *Quality of Medical Care Provided by Family Physicians as Related to Their Education, Training and Methods of Practice*. New York: Health Insurance Plan of Greater New York, 1958.
10. Lyle CB, Citron DS, Sugg WC, Williams OD. Cost of Medical Care in a Practice of Internal Medicine: A Study in a Group of Seven Internists. *Annals of Internal Medicine* 1974;81:1-6.
11. Wright DD, Kane RL, Snell GF, Wolley FR. Costs and Outcomes for Different Primary Care Providers. *Journal of the American Medical Association* 1977;238:46-50.
12. Eisenberg JM, Nicklin D. Use of Diagnostic Services by Physicians in Community Practice. *Medical Care* 1981;19:297-309.
13. Schroeder SA, Schlifftman A, Piemme TE. Variation Among Physicians in the Use of Laboratory Tests: Relation to Quality of Care. *Medical Care* 1974;12:709-13.
14. Freeborn DK, Baer D, Greenlick MR, Bailey JW. Determinants of Medical Care Utilization: Physicians' Use of Laboratory Services. *American Journal of Public Health* 1972;62:846-53.
15. Weeks HA. *Tightness of Organization Related to Patterns of Patient Care*. Ann Arbor: University of Michigan, School of Public Health, Bureau of Public Health Economics, n.d.; Sparling JF. Measuring Medical Care Quality: A Comparative Study. *Hospitals* 36 1962;16:62-8; 36 1962;1:56-7,60-61.
16. Doyle JC. Unnecessary Hysterectomies: Study of 6,248 Operations in Thirty-Five Hospitals During 1948. *Journal of the American Medical Association* 1953;51:360-5.
17. Bunker JP, Forrest WH, Jr, Mosteller F, Vandam LD (eds.). *The National Halothane Study: A Study of the Possible Association Between Halothane Anesthesia and Postoperative Hepatic Necrosis*. Bethesda MD. National Institutes of Health, National Institute of General Medical Sciences, 1969; Stanford Center for Health Care Research. Comparison of Hospitals With Regard to Outcomes of Surgery. *Health Services Research* 1976;11:112-27; Scott WR, Forrest WH, Jr, Brown BW. Hospital Structure and Postoperative Mortality and Morbidity. in *Organizational Research in Hospitals*, ed. Shortell SM, Brown M. Chicago: Blue Cross Association, 1976; Scott WR, Flood AB, Ewy W. Organizational Determinants of Services, Quality and Cost of Care in Hospitals. *Milbank Memorial Fund Quarterly: Health and Society* 1979;57:234-64; Flood AB, Scott WR. Professional Power and Professional Effectiveness: The Power of the Surgical Staff and the Quality of Surgical Care in Hospitals. *Journal of Health and Social Behavior* 1978;19:240-54; Flood AB, Ewy W, Scott WR, Forrest WH, Jr, Brown BW. The Relationship Between Intensity and Duration of Medical Services and Outcomes for Hospitalized Patients. *Medical Care* 1979;17:1088-102; Flood AB, Scott WR, Wayne E. Does Practice Make Perfect? 1: The Relation Between Hospital Volume and Outcomes for Selected Diagnostic Categories. *Medical Care* 1984;22:98-114; Flood AB, Scott WR, Wayne E. Does Practice Make Perfect? 2: The Relation Between Volume and Outcomes and Other Hospital Characteristics," *Medical Care* 1984;22:115-25.
18. New York Academy of Medicine, Committee on Public Health Relations. *Maternal Mortality in New York City: A Study of All Puerperal Deaths 1930-1932*. New York: Oxford University Press, for the Commonwealth Fund, 1933.
19. Morehead MA, et al. *A Study of the Quality of Hospital Care Secured by a Sample of Teamster Family Members in New York City*. New York: Columbia University, School of Public Health and Administrative Medicine, 1964; Rosenfeld LS. Quality of Medical Care in Hospitals. *American Journal of Public Health* 1957;47:856-65; Rhee SO. Relative Influence of Specialty Status, Organization of Office Care and Organization of Hospital Care on the Quality of Medical Care: A Multivariate Analysis. Doctoral dissertation, University of Michigan, 1975; Rhee SO. Factors Determining the Quality of Physician Performance in Patient Care. *Medical Care* 1976;14:733-50.
20. Payne BC, Lyons TF, Neuhaus E, Kolton M, Dwarshius L. *Method for Evaluating and Improving Ambulatory Care*. Ann Arbor, MI: Health Services Research Center, 1978; Morehead MA, Donaldson RS, Seravelli MR. Comparisons Between OEO Neighborhood Health Centers and Other Health Care Providers of Ratings of the Quality of Health Care. *American Journal of Public Health* 1971; 61:1294-306.
21. Ray WA, Federspiel CE, Schaffner W. Prescribing of Chloramphenicol in Ambulatory Practice: An Epidemiologic Study Among Tennessee Medicaid Recipients. *Annals of Internal Medicine* 1976;84:266-70.
22. Brook RH, Williams KN. Evaluation of the New Mexico Peer Review System, 1971 to 1973. *Medical Care* 1976;14:112-22.
23. Gittelsohn and Wennberg (nota 4).
24. Nickerson RJ, Colton T, Peterson OL, Bloom BS, Hauck WW, Jr. Doctors Who Perform Operations: A Study on In-Hospital Surgery in Four Diverse Geographic Areas. *New England Journal of Medicine* 1976;295:921-6,982-9.
25. Peterson, et al. (nota 7); Morehead (nota 9).
26. Morehead (nota 9).
27. Ibid, Peterson, et al. (nota 7).
28. Ibid.
29. Childs AW, Hunter ED. Non-Medical Factors Influencing Use of Diagnostic X-Ray by Physicians. *Medical Care* 1972;10: 323-5.
30. Brook and Williams (nota 22).
31. Payne BC, Lyons TF, Dwarshius L, Kolton M, Morris W. *The Quality of Medical Care: Evaluation and Improvement*. Chicago: Hospital Research and Educational Trust, 1976; Morehead, et al. (nota 19).
32. Ibid.
33. Riedel RL, Riedel DC. *Practice and Performance: An Assessment of Ambulatory Care*. Ann Arbor, MI: Health Administration Press, 1979; Payne, et al. (nota 20).
34. Kohl SG. *Perinatal Mortality in New York City: Responsible Factors*. Cambridge, MA: Harvard University Press, for the Commonwealth Fund, 1955.
35. Rhee SO, Luke RD, Lyons TF, Payne BC. Domain of Practice and the Quality of Physician Performance. *Medical Care* 1981;19: 14-23; Rhee. Relative Influence of Specialty Status (nota 19); Payne, et al. (nota 20); Payne, et al. (nota 31).
36. New York Academy of Medicine (nota 18); Morehead, et al. (nota 20).
37. Rhee. Relative Influence of Specialty Status (nota 19); Rhee. Factors Determining the Quality of Physician Performance (nota 19).

38. Fleming GV. Hospital Structure and Consumer Satisfaction. *Health Services Research* 1981;16:43-63.
39. Luft HS, Bunker JP, Enthoven AC. Should Operations Be Regionalized? The Empirical Relation Between Surgical Volume and Mortality. *New England Journal of Medicine* 1979;301:1364-9; Flood, *et al.* (nota 17), parts 1 and 2.
40. New York Academy of Medicine (nota 18).
41. American College of Obstetricians and Gynecologists, Committee on Maternal Health. *National Study of Maternity Care: Survey of Obstetric Practice and Associated Services in Hospitals in the United States*. Chicago: American College of Obstetricians and Gynecologists, 1970.
42. Eisenberg and Nicklin (nota 12).
43. Childs and Hunter (nota 29).
44. Brook and Williams (nota 22).
45. Peterson, *et al.* (nota 7); Clute (nota 7).
46. Morehead (nota 9).
47. Payne, *et al.* (nota 20); Payne, *et al.* (nota 31).
48. Rhee, "Factors Determining the Quality of Physician Performance" (nota 19).
49. Riedel and Riedel (nota 33).
50. Peterson, *et al.* (nota 7).
51. *Ibid.*
52. Clute (nota 7).
53. Morehead (nota 9).
54. *Ibid.*
55. Payne, *et al.* (nota 20).
56. Hulka, *et al.* (nota 8).
57. Childs and Hunter (nota 29).
58. United Steelworkers of America, Public Relations Department, Special Study on the Medical Care Program for Steelworkers and Their Families (Pittsburgh: USA, 1960).
59. Sparling (nota 15).
60. Fitzpatrick RB, Riedel DC, Payne BC. Appropriateness of Admission and Length of Stay. In: *Hospital and Medical Economics: A Study of Population, Services, Costs, Methods of Payment, and Controls*. vol. 1, by: McNerney WJ, *et al.* Chicago: Hospital Research and Educational Trust, 1962.
61. Donabedian A. An Evaluation of Prepaid Group Practice. *Inquiry* 1969;6:3-27; Roemer MI, Shonick W. HMO Performance: The Recent Evidence. *Milbank Memorial Fund Quarterly: Health and Society* 1973;51:271-317; Luft HS. Assessing the Evidence on HMO Performance. *Milbank Memorial Fund Quarterly: Health and Society* 1980;58:501-36.
62. Riedel and Riedel (nota 33).
63. Peterson, *et al.* (nota 7).
64. Clute (nota 7).
65. Payne, *et al.* (nota 31).
66. *Ibid.*
67. Rhee, "Relative Influence of Specialty Status" (nota 19); Rhee, "Factors Determining the Quality of Physician Performance" (nota 19).
68. Morehead (nota 9).
69. *Ibid.*
70. Peterson, *et al.* (nota 7).
71. Gordis L. Effectiveness of Comprehensive-Care Programs in Preventing Rheumatic Fever. *New England Journal of Medicine* 1973; 289:331-5.
72. Shapiro S, Weiner L, Densen PM. Comparison of Prematurity and Perinatal Mortality in a General Population and in the Population of a Prepaid Group Practice Medical Care Plan. *American Journal of Public Health* 1985;48:170-87.
73. Quick JD, Greenlick MR, Roghmann KJ. Prenatal Care and Pregnancy Outcome in an HMO and General Population: A Multivariate Cohort Analysis. *American Journal of Public Health* 1981;71:381-90.
74. Morehead, *et al.* (nota 19).
75. Sparling (nota 15).
76. Fine J, Morehead MA. Study of Peer Review of Inhospital Patient Care. *New York State Journal of Medicine* 1971;71:1963-73; Morehead, *et al.* (nota 19); Rhee. Relative Influence of Specialty Status (nota 19); Rhee. Factors Determining the Quality of Physician Performance (nota 19).
77. Lipworth L, Lee JAH, Morris JN. Case-Fatality in Teaching and Non-Teaching Hospitals 1956-59. *Medical Care* 1963;1:71-6.
78. Kohl (nota 34).
79. American College of Obstetricians and Gynecologists (nota 41).
80. Stanford Center for Health Care Research, Study of Institutional Differences in Postoperative Mortality. Springfield, VA: U.S. Department of Commerce, National Technical Information Service, 1974); Scott, *et al.* Hospital Structure (nota 17).
81. Luft, *et al.* (nota 39).
82. Roemer MI, Moustafa AT, Hopkins CE. A Proposed Hospital Quality Index: Hospital Death Rates Adjusted for Case Severity. *Health Services Research* 1968;3:96-118.
83. Goss MEW, Reed JI. Evaluating the Quality of Hospital Care Through Severity-Adjusted Death Rates: Some Pitfalls. *Medical Care* 1974;12:202-13.
84. Kohl (nota 34).
85. Morehead, *et al.* (nota 19).
86. Stanford Center for Health Care Research (nota 80).
87. Roemer, *et al.* (nota 82).
88. Morehead, *et al.* (nota 19).
89. Neuhauser D. *The Relationship Between Administrative Activities and Hospital Performance*. Research Series 28 (Chicago: University of Chicago, Center for Health Administrative Studies, 1971).
90. Morehead, *et al.* (nota 19).
91. Rhee. Relative Influence of Specialty Status (nota 19).

92. Georgopoulos BS, Mann FC. The Community General Hospital New York: Macmillan Co., 1962; Shortell SM, Becker SW, Neuhauser D. The Effects of Managerial Practices on Hospital Efficiency and Quality of Care. In *Organizational Research in Hospitals*, ed. Shortell SM, Brown M. Chicago: Blue Cross Association, 1976. Scott, *et al.* Hospital Structure (nota 17); Scott, *et al.* Organizational Determinants (nota 17); Flood and Scott (nota 17); Neuhauser (note 89).
93. Scott, *et al.* Hospital Structure" (nota 17); Scott, *et al.* Organizational Determinants (nota 17); Shortell, *et al.* (nota 92).
94. Scott, *et al.* Hospital Structure" (nota 17); Flood and Scott (nota 17).
95. Flood and Scott (nota 17).
96. Scott, *et al.* Organizational Determinants" (nota 17).
97. Wyszewianski L, Donabedian A. Equity in the Distribution of Quality of Care. *Medical Care* 1981;19:28-56.
98. Brook RH. *Quality of Care Assessment: A Comparison of Five Methods of Peer Review*. Washington, DC: Department of Health, Education, and Welfare, Public Health Service, Health Resources Administration, Bureau of Health Services Research and Evaluation, 1973.
99. Ray, *et al.* (nota 21).
100. Bombardier C, Fuchs VR, Lillard LA, Warner KE. Socioeconomic Factors Affecting the Utilization of Surgical Operations. *New England Journal of Medicine* 1977;297:699-705.
101. Lembcke (nota 4).
102. Lyons TF, Payne BC. The Quality of Physicians' Health Care Performance: A Comparison Against Optimal Criteria for Treatment of the Elderly and Younger Adults in Community Hospitals. *Journal of the American Medical Association* 1974;227:925-8; Lyons TF, Payne BC. Quality of Physicians' Office-Care Performance: Different for Elderly Than for Younger Adults?. *Journal of the American Medical Association* 1974;229:1621-2.
103. Janzen E. Quality Nursing Care Assurance: Initial Survey. Paper presented at the American Public Health Association annual meeting, New Orleans, Oct. 23, 1974.
104. Lichtner S, Pflanz M. Appendectomy in the Federal Republic of Germany: Epidemiology and Medical Care Patterns. *Medical Care* 1971;9:311-30.
105. Rhee SO, Lyons TF, Payne BC. Patient Race and Physician Performances: Quality of Medical Care, Hospital Admissions and Hospital Stays. *Medical Care* 1979;17:737-47.
106. Bombardier, *et al.* (nota 100).
107. Janzen (nota 103).
108. Egbert LD, Rothman IL. Relation Between the Race and Economic Status of Patients and Who Performs Their Surgery. *New England Journal of Medicine* 1977;297:90-1.
109. Morehead, *et al.* (nota 19).
110. Kohl (nota 34).
111. Bunker JP, Brown BW, Jr. The Physician-Patient as an Informed Consumer of Surgical Services. *New England Journal of Medicine* 1974;290:1051-5.
112. Ibid.
113. Lembcke PA. Medical Auditing by Scientific Methods: Illustrated by Major Female Pelvic Surgery. *Journal of the American Medical Association* 1956;162:646-55; McCarthy EG, Finkel ML, Ruchlin HS. *Second Opinion Elective Surgery*. Boston: Auburn House, 1981; Doyle (nota 16).
114. Lichtner and Pflanz (nota 104).
115. Howie JGR. The Place of Appendectomy in the Treatment of Young Adult Patients With Possible Appendicitis. *Lancet* 1968, 1365-7.
116. Fitzpatrick, *et al.* (nota 60).
117. Sparling (nota 15).
118. Flood and Scott (nota 17).
119. New York Academy of Medicine (nota 18).
120. Kohl (nota 34).
121. Shapiro, *et al.* (nota 72).

Gestión del conocimiento y calidad asistencial

Donabedian A. La investigación sobre la calidad de la atención médica.

Salud Pública Mex 1986;28:324-7

Avedis Donabedian estuvo en México hacia finales de los años setenta invitado por un distinguido médico mexicano, por naturaleza innovador, el Dr. José Laguna. Fue en aquella ocasión cuando Julio Frenk conoció a Avedis y seguramente fue ese encuentro una de las marcas indelebles en la vida profesional de Julio. Poco tiempo después iría a estudiar a la Universidad de Michigan con quien sería muy pronto su dilecto maestro y amigo de siempre y se inició una verdadera e impecable sinergia intelectual cuyos efectos han sido patentes en este país hasta la fecha. Avedis falleció unos cuantos días antes del nombramiento de Julio Frenk como Secretario de Salud de México y casi veinticinco años después de aquel primer encuentro.

Este artículo adquiere un interesante significado en este contexto pues fue leído por el propio Avedis en julio de 1985 durante su primera visita como miembro del Comité Consultivo del Centro de Investigaciones en Salud Pública (CISP) de la Secretaría de Salud de México, fundado apenas once meses antes de esa ocasión, en Agosto de 1984, por el Dr. Guillermo Soberón, entonces Secretario de Salud. El primer director de ese organismo era precisamente el Dr. Julio Frenk.

Fue esa época, y bajo la influencia del Profesor Donabedian, en la cual la investigación sobre la calidad de la atención médica empezó a adquirir presencia en este país. José Luis Bobadilla, lamentablemente fallecido en 1996, entonces Subdirector de Investigación en Necesidades de Salud del CISP, realizaba su tesis doctoral sobre la evaluación de la calidad de la atención perinatal. Diez meses antes de la conferencia del Dr. Donabedian yo había iniciado los primeros círculos de calidad en un hospital mexicano. Para marzo de 1985 y a invitación de Julio, me

incorporaría al CISP como Subdirector de Investigación en Sistemas de Salud para continuar este esfuerzo como un proyecto de investigación-acción. En consecuencia de todo lo anterior, se estableció en el CISP una línea de investigación sobre calidad de la atención médica, acontecimiento éste sin precedentes en la investigación en salud pública en México. Donabedian fue, sin duda, el inspirador, maestro y elegante y bondadoso pero firmemente crítico que nos hizo a todos madurar. Por ello, el momento en el que se inscribe este artículo es, sin duda y por muchas razones, un parteaguas en la historia de la salud pública de este país de la que Avedis fue y sigue siendo protagonista.

En este artículo, como en todos sus escritos, Avedis despliega su poderoso rigor académico y una claridad intelectual que nos permite trazar precisos mapas mentales del conocimiento aderezados con una envidiable parsimonia que, con gran tacto, exigía de los demás. Además, es muy probable que sea éste uno de los primeros artículos en los que define el concepto de "epidemiología de la calidad". En éste clasifica también el nacimiento de lo que hoy se denomina "medicina basada en evidencias". Como siempre, Avedis Donabedian se caracterizó por su profundo humanismo del cual también dejó huella en este artículo al defender la importancia de la calidad sobre los costes y la trascendencia del ser humano por sobre todo lo demás: "el secreto de la calidad está en el corazón y el alma de los hombres".

Enrique Ruelas

*Subsecretario de Innovación y Calidad
Secretaría de Salud. México*

La investigación sobre la calidad de la atención médica*

Avedis Donabedian

Resumen

Se distinguen los dos grandes dominios de la investigación sobre la calidad de la atención médica, uno clínico y el otro socio-organizativo. En seguida se describen algunas áreas específicas, principalmente desde el punto de vista socio-organizativo. La primera tiene que ver con el concepto de calidad. Después se definen los medios para alcanzar la mejor calidad de atención: estrategias económicas y aprovechamiento de innovaciones técnicas. Resulta necesario medir las organizaciones, la "calidad de la vida" y la precisión de las fuentes de información. Así se llega a la tercera área: epidemiología de la calidad, que describe la distribución de la calidad entre quienes proporcionan la atención y sus beneficiarios. La cuarta área de investigación es el proyecto, la implantación y la evaluación de sistemas de monitoreo dentro y fuera de la organización. Por último, la quinta área de investigación, será el estudio de comportamiento humano, pues el secreto de la calidad está en los hombres.

El tema es tan amplio que apenas sabría por dónde empezar o qué decir. Casi todos estaríamos de acuerdo en que el propósito de los sistemas de atención a la salud, en su núcleo y a través de sus numerosas partes, es proporcionar el más alto nivel de calidad al menor coste, de la manera más equitativa y al mayor número de personas. Si esto es así, la investigación de la calidad comprende el estudio de todas las organizaciones y actividades que persiguen este noble objetivo.

No obstante, esta definición es tan general que casi no resulta de valor. Para saber de qué elementos se compone, debemos separarla en todas sus partes.

Empezaré por distinguir (como, en efecto, lo hace nuestro programa) dos dominios principales: a uno lo llamaré "investigación clínica" y al otro "investigación socio-organizativa". Con ello, y esto hay que dejarlo muy claro, no trato de establecer límites artificiales ni señalar territorios abiertos para unos y prohibidos para otros. Quiero usar mi clasificación tan sólo para establecer ciertas distinciones de concepto, de política y de

Presentado en el Simposio sobre "La Calidad de la Atención Médica" organizado por la Secretaría de Salud, El Colegio Nacional, la Academia Nacional de Medicina; con el patrocinio del Consejo Nacional de Ciencia y Tecnología y la Agencia para el Desarrollo Internacional, el 5 de julio de 1985.

Summary

There are two principal areas of research concerning the quality of medical care, one refers to the clinical aspects and the other to social-organizational aspects. The article describes both of these areas of research, emphasizing the social-organizational aspects. First the concept of quality care is defined. Then the methods for improving the quality are discussed, including economic strategies and the implementation of new medical techniques. As part of this process, indicators must be developed for organization variables, quality of life, and the reliability of the sources of information. A third aspect refers to the epidemiology of quality care, that describes the distribution of quality among those groups who provide the care and their patient-clients. Another aspect of research deals with the planning, implementation, and the evaluation of monitoring systems both within and outside of the organizations. The last part of the article concerns the need for further study of human behaviour because the degree of quality in any system ultimately depends on human factors.

responsabilidad primaria. Para ayudarme en tal empresa, echaré mano de la conocida triada de estructura, proceso y resultado.

Como se sabe, "estructura" se refiere a las características más establecidas e invariantes de cómo se establece y maneja el sistema de atención a la salud. "Proceso" se refiere fundamentalmente a lo que hacen los médicos, desde el punto de vista técnico, en la conducción de la atención a la salud, y también en cuanto a su interacción con los pacientes. "Resultados" son las consecuencias de la atención (o de la falta de ella) en la salud de quienes la reciben (o dejan de recibirla).

La utilidad de esta formulación se basa por completo en la observación de que las características estructurales influyen en el proceso de la atención, y que éste lo hace en la salud. Así, en el corazón mismo de la investigación de la calidad está el esclarecimiento de la relación entre el proceso y el resultado, y entre la estructura y el proceso.

La investigación clínica se ocupa de elucidar y fortalecer la relación entre el proceso de la atención y sus resultados esperados. Se preocupa, primordialmente, de la eficacia de la atención en condiciones óptimas o, cuando menos, en condiciones razonablemente favorables.

La investigación clínica tiene tres clases de instrumentos:

Primero, puede tratar de reunir y formalizar el conocimiento existente para identificar las estrategias más eficaces

de atención. El análisis formal de las decisiones resulta una técnica útil en esta tarea.

Segundo, la investigación clínica puede usar los instrumentos de la epidemiología para observar y evaluar relaciones entre procesos y resultados específicos. En condiciones óptimas, saca total provecho de los llamados "experimentos naturales".

El tercero y más definitivo instrumento de la investigación clínica es el propio ensayo clínico, preparado como experimento auténtico y en condiciones cuidadosamente controladas.

Por lo que hace a sus términos de referencia y a sus criterios, la investigación clínica puede considerarse como progresiva, se aproxima cada vez más a la perspectiva socio-organizativa; de suerte que la distinción entre los dos dominios empieza a hacerse borrosa.

En su estado original y tradicional la investigación clínica se ocupaba primordialmente del efecto de la atención médica en la salud física y fisiológica de cada paciente, según la evaluación de los propios investigadores y sin considerar costes. En tal virtud, constituye la base de un tipo de práctica clínica todavía legítimo y que goza de alta estima, pero que cada vez recibe mayor número de ataques.

Si la investigación clínica ha de responder más a las exigencias sociales del momento, tendrá que adoptar una perspectiva de la salud que comprenda los aspectos psicológico y social. Así, tendrá en cuenta el concepto que los pacientes informados se hacen de los diversos resultados de las distintas estrategias de atención. Incluirá el estudio de los estilos de interacción entre los pacientes y los médicos como parte integral de la "tecnología" de la atención. Lo que es más importante, atenderá a los costes, tanto directos como indirectos, en tanto que tendrá criterio clave para evaluar las diferentes formas de atención.

La investigación clínica de este nuevo tipo no sólo es posible, sino necesaria. Es la base de la nueva ciencia de la medicina, la fuente de la cual han de emerger necesariamente nuestras nuevas especificaciones de calidad. Es, asimismo, una ciencia tal fiel a sus compromisos morales como cualquiera que hayamos conocido, pero también está expuesta a una terrible tentación: la de asignar valor monetario a la vida y bienestar del hombre o, en otras palabras, a regatear y transar entre la calidad y el coste. En mi opinión, la profesión médica, en general, y los investigadores clínicos, en particular, deben negarse a juzgar cuáles deberían ser las condiciones de negociación entre la calidad y el coste. De la investigación clínica sólo puede pedirse que proporcione la información más precisa y completa sobre la relación entre el proceso de la atención y sus resultados. Que otros se encarguen de lo que podemos o no podemos pagar.

Permítanme ahora pasar al estudio de la relación entre estructura y proceso y, a través de éste, al resultado.

Es claro que sin las contribuciones de la investigación clínica no podríamos saber qué procesos de atención son buenos y cuáles peores o mejores. Pero, dada esa información, podemos estudiar gran número de problemas que, tomados en conjunto, constituyen el dominio de la investigación socio-organizativa en torno a la calidad. El propósito de esta investigación es determinar la calidad de la atención que proporcionan los sistemas o programas existentes, así como idear y sostener a prueba formas en que pueda mejorarse tal calidad.

Los métodos de esta investigación son, en esencia, semejantes a los ya descritos; a saber: la sistematización y la interpretación de la información de que se dispone, la observación epidemiológica del desempeño de los modos contrastantes de organización, y la experimentación controlada con diferentes planes de organización e instrumentos administrativos. Los criterios de calidad se derivan, como hemos dicho, de la investigación clínica, pero ahora se reformulan para que puedan ser aplicados ya no a individuos, sino al menos a grupos de pacientes y, con mayor propiedad, a comunidades y poblaciones enteras. Así pues, llegamos a ocuparnos de la distribución equitativa de la calidad y de los costes y beneficios sociales de la atención.

Una vez distinguidos estos dos grandes dominios de la investigación sobre la calidad, uno clínico y el otro socio-organizativo, me gustaría describir, en seguida, algunas áreas más específicas que considero particularmente importantes y potencialmente fructíferas si se consideran no exclusivas, sino principalmente desde el punto de vista socio-organizativo.

La primera área de investigación es fundamental a todo investigador de la calidad, ya sea clínico o socio-organizativo. Tiene que ver con elucidar el concepto de la calidad misma. En particular, nos gustaría saber cómo conciben y definen la calidad los participantes clave en la atención a la salud; cómo la perciben los pacientes, o los médicos, o los administradores, o quienes formulan las políticas; cuáles son los puntos de semejanza o discrepancia, y cómo pueden conciliarse los diversos puntos de vista.

Estas investigaciones se refieren tanto a los objetivos de la atención como a los medios para conseguirlos.

La elucidación de objetivos no conduce a indagaciones casi filosóficas del significado mismo de la salud.

El estudio de los medios trasciende la especificación de las intervenciones clínicas; abarca cuestiones morales de gran complejidad. Por ejemplo, ¿en qué medida los individuos son libres al hacer sus elecciones y cuándo?, si acaso, puede usarse el engaño o la coerción para alcanzar fines sociales.

Entre los medios, el precio monetario, como ya lo he insinuado, asume un papel importante.

Ya he dicho que la nueva ciencia clínica de la medicina considerará el coste monetario como uno de los elementos para la evaluación de las estrategias posibles de atención. Una vez identificadas las estrategias más económicas, necesitamos encontrar en qué medida, en qué lugares, en qué formas y por qué razones la práctica actual no alcanza el ideal de economía. Las sumas de dinero destinadas a corregir los excesos actuales harían menos necesario sacrificar las posibles mejoras en salud, en vista de que no podemos satisfacer sus costes. Pero también debemos recordar que en muchos casos la evaluación de la práctica actual requeriría de gastos adicionales para proporcionar niveles más altos de calidad. Esta posibilidad hace de la práctica de una medicina más económica, un objetivo aún más apremiante.

Debemos estar particularmente alertas en cuanto a juzgar las innovaciones técnicas no sólo por sus costes y logros directos, sino también por todos los demás que, en virtud de su coste elevado, nos impiden hacer.

Otra área de investigación de la calidad tiene que ver con la medición. Medir las características más obvias de las organi-

zaciones parece ser bastante fácil, pero hay otras características de gran sutileza que son difíciles de definir y medir. Pensemos, por ejemplo, en el poder, la influencia, la autonomía, la subordinación, la coordinación, la continuidad, la descentralización, la motivación, etc. Con todo, estas propiedades pueden más fácilmente tener efectos importantes en la conducta de las organizaciones y, por tanto, en la calidad de la atención.

En el otro extremo de nuestro tránsito gradual de la estructura al proceso y de éste al resultado, nos enfrentamos a la tarea de medir las muchas facetas de la salud y, a la larga, de hallar la manera de reunir las mediante alguna medida coherente de la "calidad de vida".

La propia medición del proceso, elemento intermedio de nuestra triada, se ha considerado generalmente situada en el núcleo de la investigación sobre la calidad. Pero a pesar del mucho trabajo realizado en esta área, queda mucho por hacer en cuanto al desarrollo de métodos para formular criterios de proceso que puedan usarse para llegar a juicios válidos y reproducibles en cuanto a la calidad. Parte de esta investigación se refiere a la integridad y precisión de las fuentes de información, principalmente en el registro médico, sobre las cuales se basan muy a menudo los juicios de calidad. Los métodos de evaluación de calidad que hemos tenido, así como los más refinados que esperamos tener en el futuro son, obviamente, instrumentos para usarse en la tercera área de investigación que quiero mencionar; a saber: el estudio de la epidemiología de la calidad.

A través del análisis epidemiológico empezamos a entender la naturaleza, magnitud y localizaciones de los problemas que encaramos y, asimismo, a reunir claves que nos indiquen las causas probables de nuestros problemas y sus posibles soluciones.

La epidemiología de la calidad es particularmente interesante porque puede describir la distribución de la calidad en dos poblaciones: la de quienes proporcionan la atención y la de los beneficiarios reales o potenciales de ésta.

La primera de estas dos distribuciones ha sido objeto de mucho estudio, siendo los principales hallazgos de éste la insuficiencia generalizada del desempeño relativo a las normas ostensibles de la práctica, aunada a grandes variaciones, en gran parte no explicadas, entre distintos médicos, entre instituciones, y aún entre áreas geográficas. En particular, sabemos poco de las maneras como la dinámica clínica y la participación del usuario, pero hemos empezado a conocerlas.

La segunda faceta de la epidemiología de la calidad, la distribución de esta calidad entre los beneficiarios potenciales y reales de la atención, generalmente se enfoca de manera indirecta, a través de estudios de acceso y utilización, por un lado, y de incapacidad, morbilidad y mortalidad, por el otro. Necesitamos estudiar más directamente la atención que reciben diferentes subgrupos de la población, ya sea de la misma fuente o de diferentes. Sospecho que nuestra renuncia a hacerlo proviene de nuestro temor a lo que pudiéramos descubrir.

Nuestro conocimiento de la epidemiología de la calidad nos lleva naturalmente a los sistemas de inteligencia epidemiológica.

A través de otra extensión, llegamos a lo que claramente es la cuarta área de investigación: el proyecto, la implantación y la evaluación de sistemas de monitorización dentro y fuera de la organización. En este caso, aplicamos mucho de lo que sabemos acerca de cómo puede medirse la calidad, y sobre dónde, en qué momento, en manos de quién y de qué manera es más probable que la calidad resulte más gravemente insuficiente. El propósito ahora, es planear un sistema de monitorización constituido de muchos elementos y en varios pasos que nos permita identificar rápida y eficientemente nuestros errores más graves. Pero el propósito de conocer nuestros errores, es tomar medidas para corregirlos o, todavía mejor, prevenirlos. Así llegamos a nuestra quinta y acaso más importante área de investigación: el estudio de la motivación del comportamiento humano. Creo que es justo decir que un sistema fundado predominantemente en buscar y castigar a quienes obran mal, probablemente será indeseable e ineficaz. Consecuentemente, algunos han buscado el secreto de la calidad en la educación, particularmente cuando se asocia con las diferencias individuales, cuando tales deficiencias se hacen manifiestas y son más susceptibles de corrección; otros han buscado el secreto de la calidad en la planeación organizativa, esperando encontrar la combinación de características estructurales y operacionales que constituyen el medio más adecuado para motivarlo mejor en cada uno de sus miembros.

Sin duda, todas estas cosas, al igual que algunas otras son importantes y necesarias, pero me parece que, a fin de cuentas el secreto de la calidad está en el corazón y el alma de los hombres. Allí, del manantial de lo mejor que hay en nosotros, sacamos toda inspiración a partir de algo o de alguien. Estamos con el oído atento a estas pequeñas pero imperiosas voces.

Relevancia de los criterios y estándares para la evaluación y mejora de la calidad

Donabedian A. Criteria and standars for quality assessment and monitoring. *Qual Rev Bull* 1986;12:99-108

En este artículo Avedis Donabedian nos ofrece una amplia y valiosa revisión sobre las características y naturaleza de los criterios y estándares empleados en la evaluación y mejora de la calidad. Define el significado de cada uno de estos términos y nos explica como a veces se superponen los conceptos que encierran de forma que los utilizamos de forma indistinta. Se adentra en una amplia descripción de cómo se pueden clasificar y de cuales son sus principales atributos en cuanto a las condiciones que deben de cumplir para ser capaces de discernir e identificar aquellos aspectos de la atención sanitaria que relacionamos con la calidad. Establece las pautas a seguir y los aspectos a considerar a la hora de formular criterios y estándares, seleccionar los mas adecuados a cada caso e integrar los resultados de su monitorización en el ciclo de mejora continua de la calidad.

La relevancia de cada uno de los conceptos y ejemplos expuestos por Donabedian en esta revisión es incuestionable. Hoy en día no entenderíamos ningún estudio de calidad que no los manejara. Se aplican en las evaluaciones que realizamos internamente en los centros sanitarios y en todas aquellas evaluaciones externas destinadas a obtener acreditaciones y certificaciones.

Durante las últimas semanas nos han llegado noticias sobre la decisión adoptada por el Ministerio de Sanidad y Consumo de dotar a todo el Sistema Nacional de Sanidad de un modelo de evaluación y mejora de la calidad común. Para ello, indudablemente, el primer paso será establecer criterios y estándares uniformes que además deberán ser consensuados por los profesionales y contar con una amplia aceptación. En este sentido cabe destacar iniciativas como la desarrollada por la Sociedad Española de Ginecología y Obstetricia que el pasado año publicó un compendio de criterios, estándares e

indicadores de calidad de la especialidad diseñados desde el mas puro estilo preconizado por Donabedian. Y tal como el nos dice, estos elementos no solo constituyen un método útil que facilita la evaluación y mejora, sino también una poderosa herramienta educativa y de difusión de determinados tipos de práctica. Así nos lleva a la reflexión de que "aquellos que controlan los criterios también controlan una buena parte del sistema sanitario".

Al diseñar estudios de evaluación y mejora de la calidad, o al explicar como hacerlo, siempre se tiene la sensación de que la metodología que empleamos es tremendamente intuitiva y que tiene mucho que ver con los procesos lógicos de pensamiento que utilizamos en nuestras vidas cotidianas y, por supuesto, con los que desarrollamos al ejercer la profesión médica. Donabedian nos señala en su trabajo como la evaluación es una constante de casi todas las actividades de los seres humanos y como en cada acto de evaluación, seamos conscientes o no de ello, utilizamos criterios mas o menos explícitos. De ahí que la utilización de criterios adquiera un sentido universal.

Por último es obligado señalar que además del carácter puramente descriptivo y el nivel de exhaustividad de esta exposición, Donabedian, como es habitual en sus escritos, nos obsequia con una serie de comentarios que van mucho mas allá de su aspecto técnico y nos invitan a reflexionar sobre la trascendencia de lo que estamos tratando. Por todo ello, cuando finalicemos su lectura, seguro que la archivaremos en esa carpeta del recuerdo en la que guardamos los trabajos inteligentes.

Victoria Ureña
Coordinadora de Calidad
Hospital Ramón y Cajal. Madrid

Criterios y estándares para evaluar y monitorizar la calidad*

Avedis Donabedian

Los criterios y estándares de calidad derivan del concepto previo de calidad al que representan. Son, por consiguiente, tan amplios o limitados y tan aceptables o cuestionables como los conceptos que les dieron vida. No obstante, dado que los criterios y estándares son encarnaciones concretas y específicas de unos conceptos originarios, representan la calidad de manera más inmediata, amenazando suplantar aquello para lo que fueron creados, aunque de manera imperfecta, únicamente para servir. Esa es la razón por la que los criterios y estándares de calidad merezcan nuestra atención más especial.

Definición de Términos

Cualquier debate sobre criterios y estándares de evaluación de calidad puede verse ensombrecido por diferencias de terminología. Voy a definir "criterios" como un componente o aspecto de la estructura de un sistema de atención sanitaria, o del proceso o de los resultados de la atención, que tienen relación con la calidad asistencial. Algunos ejemplos son el porcentaje enfermeros/as-camas en un hospital, la proporción de niños con dolor de garganta tratados con antibióticos y el nivel de presión arterial en los hipertensos. Un "estándar" es la especificación precisa y cuantitativa del estado de un criterio, que representa la calidad en cierto grado. Por ejemplo, un estándar podría especificar que debe haber un mínimo de un enfermero/a por cada dos camas de hospital en la unidad de cuidados intensivos, que un porcentaje inferior al 5% de los niños atendidos por un médico de familia de debería tratarse con antibióticos, o que un porcentaje no inferior al 70% de jóvenes hipertensos debería tener una presión diastólica de 90 en un plazo inferior a un año después de iniciado el tratamiento.

Un tercer término, "norma", aparece con frecuencia en la literatura sobre evaluación. A veces, tiene el significado de estándar según lo he empleado anteriormente. Otras veces, tiene el significado de un tipo de estándar especial, derivado de la observación de la práctica y se expresa según la edad o un rango específico. Yo prefiero utilizar "norma" como un concepto general sobre las cualidades positivas de un criterio, por ejemplo, la existencia de un número suficiente de enfermeros/as, o el empleo de antibióticos únicamente en caso necesario.

La mayoría de las veces, cuando hablamos de criterios, nos referimos a un compendio de lo que he descrito anterior-

mente como criterios, estándares y normas. Desgraciadamente, no tenemos una palabra para definir esta tríada. Por ello, en este artículo utilizaremos "criterios" de manera genérica para todo.

Características descriptivas

La Figura 1 detalla algunos de los modos de clasificación y definición de criterios de calidad.

Enfoque

Según el enfoque utilizado para evaluar la calidad de la asistencia, los criterios se pueden clasificar como pertenecientes a la estructura, proceso o resultado. La estructura describe las características físicas, de organización y otros rasgos del sistema asistencial y de su entorno. El proceso es lo que se hace para tratar a los pacientes. El resultado es lo que se consigue, normalmente una mejora en la salud, aunque también en la actitud, los conocimientos y en la conducta que derivará en una salud futura.

Fuente

Según su fuente, los criterios se pueden clasificar como "de base normativa" o "de base empírica". Los primeros representan lo que las fuentes mejor informadas dicen que debería hacer o conseguirse. Los últimos representan aquello que se ha realizado o conseguido, valorado tanto como medias, como a través de la práctica de médicos o instituciones seleccionadas, que podrían utilizarse como estándar de comparación. La ventaja de los estándares empíricos es que se puede demostrar que es posible cumplirlos. Su desventaja es que pueden quedarse por debajo de lo que se puede conseguir.

Referente

Los criterios pueden clasificarse también según sus referentes, es decir, las condiciones o situaciones a que pertenecen. En la mayoría de los casos, este es un diagnóstico clínico: hipertensión, infección urinaria, diabetes, etc. Las categorías de diagnósticos pueden necesitar subdividirse en subgrupos más homogéneos en lo referente al manejo y a los resultados

alcanzables. El objetivo es conseguir una correspondencia más estrecha entre los criterios y lo que se juzga que hay que cumplir. Para evitar una falta mayor de correlación, deberían existir criterios que verificaran la exactitud de los diagnósticos. No obstante, estas salvaguardas no se mantienen cuando el referente sea un diagnóstico uno no puede identificar fácilmente aquellos casos que podrían haberse diagnosticado y se omitieron.

El referente puede ser una condición o problema representado, a menudo, por un conjunto de signos o síntomas: dolor abdominal, cefalea, coma, etc. Entonces, la capacidad diagnóstica se puede evaluar de manera más directa y se identifican mejor los diagnósticos omitidos. El referente también puede ser una intervención quirúrgica, un procedimiento diagnóstico o un determinado tipo de terapia. Incluso puede referirse a un lugar en el que se produce la atención como el hospital, residencia, o clínica.

Cuando el referente es un centro, los criterios son criterios de "nivel de asistencia", que especifican los servicios que debe recibir un paciente para justificar el ingreso o la permanencia en el hospital o en otro centro (Figura 2). Como precaución, puede necesitarse información adicional sobre la gravedad de la enfermedad. Por otra parte, la estancia se puede justificar, simplemente facilitando los servicios requeridos, aunque no sean estrictamente necesarios.

Criterios implícitos y explícitos

En otro método, los criterios se pueden clasificar, según el grado de especificación previa, como "implícitos" o "explícitos". Cuando se emplean criterios completamente implícitos, se solicita a un clínico experto que evalúe la calidad de la atención utilizando su propio juicio. No existen otras directrices. El resultado depende del juicio, cuidado e interés de los que juzgan. Si se dispone del número necesario de expertos podrá evaluarse la totalidad de los campos de la asistencia. No obstante, este método es muy laborioso, caro y está sometido a la variabilidad de opinión de los que juzgan. Para mejorar la fiabilidad, se puede pedir que los juicios para evaluar un caso determinado los emita más de un juez, actuando cada uno independientemente de los demás. Posteriormente, los jueces podrán o no, reunirse para debatir y ajustar sus diferencias.

Contrariamente, los criterios explícitos se especifican por escrito antes de realizar cualquier evaluación. La preparación de los criterios exige una experiencia y un juicio considerable así como una considerable cantidad de tiempo. Pero una vez elaborados los criterios, la evaluación de los casos individuales no requiere de mucho juicio; puede llevarse a cabo de manera precisa, con diligencia y con unos costes relativamente bajos por parte de personal administrativo o semiprofesional. Por estos motivos, los criterios explícitos han llegado a dominar el campo de la evaluación de la calidad. Sin embargo, están sometidos a considerables inconvenientes. Por definición, los criterios explícitos se pueden aplicar únicamente a aquellos casos para los que han sido elaborados. Fundamentalmente, no pueden adaptarse completamente a todas las variaciones individuales, que permanecerán en cualquier categoría de casos incluso después de que la categoría se haya dividido en subgrupos relativamente homogéneos.

Una respuesta a los puntos fuertes y débiles de los criterios explícitos sería utilizarlos conjuntamente con juicios basados en criterios implícitos. Los criterios explícitos más simples pueden utilizarse, en primer lugar, para realizar una selección rápida de los casos en dos grupos: los probablemente bien manejados y los mal manejados. La totalidad o una muestra de estos últimos (y posiblemente una muestra también más reducida de los primeros) se evalúan más detalladamente, quizás utilizando criterios explícitos más detallados, con el añadido del juicio de un experto.

Habría que destacar que los criterios totalmente implícitos y totalmente explícitos, son únicamente dos extremos de una continuidad, con muchas variantes entre medio. Por ejemplo, los evaluadores que utilizan criterios implícitos se pueden guiar por instrucciones bastante detalladas acerca de qué y cómo evaluar. También se les puede preguntar que expliquen por qué han desaprobado determinados casos. De este modo, los criterios implícitos se especifican, de hecho, con posterioridad, convirtiéndose en explícitos y sujetos a revisión y debate.

Formato

Los criterios también se pueden clasificar según el formato como "lineales", "con ramificaciones parciales" y "completamente ramificados". Los criterios lineales, también conocidos vulgarmente como criterios de la "lista de lavandería", son,

Figura 1. Los criterios de calidad pueden clasificarse de acuerdo a varias características

Algunas características descriptivas de los criterios

- A. Basados en un enfoque de la evaluación
 - 1. Perteneciente a la estructura
 - 2. Perteneciente al proceso
 - 3. Perteneciente a los resultados
- B. Basados en las fuentes
 - 1. Derivados de normas
 - 2. Derivados de forma empírica
- C. Basados en referentes
 - 1. Referidos a diagnósticos
 - 2. Referidos a problemas o condición
 - 3. Referidos a procedimientos
 - 4. Referidos a centros (criterios de "nivel de asistencia")
- D. Basados en el grado de especificación
 - 1. Implícitos
 - a. Totalmente implícitos
 - b. Implícitos
 - c. Implícitos con especificaciones posteriores
 - 2. Explícitos
- E. Basados en un formato
 - 1. Lineales
 - 2. Parcialmente ramificados
 - 3. Completamente ramificados ("mapas de criterios")
- F. Basados en la ponderación de los apartados
 - 1. De igual ponderación
 - 2. De distinta ponderación

Figura 2. Un ejemplo de los criterios del nivel de atención

Criterios de Adecuación (Incluyendo los criterios del Nivel de Atención)

Servicios Médicos

- A. Intervención en quirófano el mismo día.
- B. Intervención en quirófano planificada para el día siguiente, requiriendo consulta o evaluación preoperatoria.
- C. Cateterización cardíaca el mismo día.
- D. Angiografía el mismo día.
- E. Biopsia de órgano interno el mismo día.
- F. Toracentesis o paracentesis el mismo día.
- G. Procedimiento diagnóstico invasivo del sistema nervioso central (p.ej. incisión lumbar, válvula cisternal, válvula ventricular, pneumoencefalografía) el mismo día.
- H. Cualquier prueba que requiera un control estricto de la dieta para establecer la duración de la misma.
- I. Tratamiento nuevo o experimental que requiere adaptar la dosis con frecuencia, bajo la supervisión directa del médico.
- J. Control médico estricto por parte de un médico un mínimo de tres veces por día (las observaciones realizadas deberán documentarse en la historia)
- K. Día postoperatorio para cualquier intervención realizada en los apartados A o C-G.

Cuidados de Enfermería

- A. Cuidados respiratorios: utilización continua o intermitente de atención respiratoria y/o terapia de inhalación (pecho PT, IPPB) un mínimo de 3 veces al día.
- B. Terapia parenteral: Uso continuo o intermitente de fluido IV con cualquier suplemento (electrolitos, proteína, medicamentos).
- C. Supervisión continua de constantes vitales cada 30 minutos como mínimo durante al menos cuatro horas.
- D. Inyecciones intramusculares o subcutáneas un mínimo de dos veces al día.
- E. Medición de entradas y salidas (balance).
- F. Heridas quirúrgicas importantes y drenaje (intubación pectoral. Tubos-T, hemovacs, drenaje Penrose).
- G. Supervisión médica por parte de enfermeros/as un mínimo de 3 veces al día, según las directrices del médico.

Factores del estado del paciente

- A. Durante las 24 horas anteriores al día de la revisión
 - 1. Incapacidad de vaciar o mover los intestinos (en las anteriores 24 horas) no imputable a trastornos neurológicos.
- B. En 48 horas antes del día de la revisión
 - 1. Transfusión por pérdida de sangre
 - 2. Fibrilación ventricular o "signos de ECG de isquemia grave, según se anotó en el curso clínico o en el informe ECG.
 - 3. Temperatura rectal mínima de 101 F (un mínimo de 100 F oral), cuando se ingresó al paciente por motivos distintos a la fiebre.
 - 4. Al menos una hora en estado inconsciente o coma.
 - 5. Estado confusional grave no derivado del abandono del alcohol.
 - 6. Trastornos hematológicos graves, neutropenia importante, anemia, trombocitopenia, leucocitosis, eritrocitosis o trombocitosis, con signos o síntomas manifiestos.
 - 7. Dificultades neurológicas progresivas graves.
- C. Catorce días antes del día de la revisión.
 - 1. Nuevo episodio de infarto de miocardio grave o accidente cerebrovascular (ictus).

Fuente: Gertman PM, Restuccia J. The appropriateness evaluation protocol: una técnica para evaluar los días de estancia hospitalaria innecesarios. *Med Care* 1981;19:69-87. Adaptado con permiso de JB Lippincott & Company

con mucho, la forma más utilizada actualmente (Figura 3). Su referente es, básicamente, una categoría de diagnóstico, con o sin más subclasificaciones. Sirven para representar aquellos elementos del manejo de un caso "típico" o "medio" que podrían representar un nivel de calidad acordado. A menudo, aunque no de modo invariable, la lista se inicia con criterios que justifican el primer diagnóstico asignado al caso en el alta. A continuación, los elementos necesarios para el manejo del "caso típico", clasificados en apartados como: historia clínica, explora-

ción física, pruebas de laboratorio, otros procedimientos diagnósticos, tratamiento y seguimiento. A menudo, para acomodar la variabilidad entre los casos, existe un apartado que describe las intervenciones que se consideran consistentes con el diagnóstico, lo que significa que no se van a criticar tanto si se realizan como si se omiten. Puede existir también una lista de complicaciones que, en caso de existir, justificarían el abandono del manejo prescrito del caso "típico" prescrito, incluyendo el tiempo de estancia en el hospital recomendado.

Figura 3. Ejemplo de criterios lineales

Lista de criterios (lineales) para colecistitis y colelitiasis

I. Indicaciones para el ingreso		2. Examen físico: sin especificar	
A. Criterios aceptables para la admisión de pacientes que aparentemente padecen una colecistitis grave.		3. Laboratorio	
1. Dolores, náuseas y vómitos		a. Análisis completo de sangre (hematocritos, Leucocitos, diferencial)	0,5
2. Frecuentes ataques de vesícula		b. Análisis de orina	0,5
3. Fiebre		c. Electrocardiograma si se tienen +50 años (en el plazo de 6 meses por regla general)	1,0
4. Ictericia		4. Radiología	
5. Masa en el cuadrante superior derecho		a. Rayos X de tórax en el plazo de un año	1,0
B. Diagnóstico de colecistitis o colelitiasis, ingreso para intervención quirúrgica.		b. Colecistografía	3,0
II. Servicios recomendados		III. Preoperatorio	
A. Colecistitis grave		A. Pacientes con historial de ataques recurrentes de vesícula pueden ingresarse para su estudio e intervención; se pueden necesitar hasta 5 días de preoperatorio.	
1. Historia clínica: Con especial referencia a:	Peso	IV. Duración probable de la estancia	
a. Tipo de dolor	3,0	A. Sin cirugía: puede recibir el alta al remitir la enfermedad actual	
b. Recurrencia	0,5	B. Colecistectomía sin complicaciones: de 5 a 10 días de postoperatorio	
c. Irradiación	0,5	C. Si también se requiere explorar el conducto y se deja un tubo T para drenaje: 5-15 días de postoperatorio.	
d. Síntomas relativos a ictericia	0,5	D. La estancia se prolonga y se hace impredecible si se realiza cualquiera de las siguientes intervenciones:	
e. Fecha de aparición	0,5	1. Duodenotomía	
2. Examen físico con referencia especial a:		2. Esfinterotomía	
a. Cuadrante superior derecho	3,0	3. Colecistotomía	
b. Masa	0,5	4. Reparación del conducto	
c. Tamaño del hígado	0,5	V. Complicaciones que pueden prolongar la estancia	
d. Ictericia	0,5	A. Infección de la herida	
e. Vigilia	0,5	B. Apertura de la herida	
3. Laboratorio		C. Piedra retenida en el conducto	
a. Análisis completo de sangre (hematocritos, Leucocitos, diferencial)	0,5	D. Flebotrombosis o tromboembolia	
b. Análisis de orina	0,5	E. Pancreatitis	
c. Electrocardiograma (si se tiene +50 años)	0,5	F. Pneumonía	
d. Suero Amylase (en alcohólicos o si la naturaleza del dolor es difusa)	1,0	G. Diabetes	
En caso de deshidratación:		H. Enfermedad cardíaca	
- sodio	0,5	I. Peritonitis biliar	
- cloro	0,5	J. Ictericia postoperatoria	
- dióxido de carbono	0,5	K. Ictericia sin explicación	
- potasio	0,5	L. Fiebre de origen desconocido	
En caso de ictericia:		VI. Indicadores para el alta	
- bilirubina	1,0	A. Sin cirugía	
- fosfata alcalina	1,0	1. Sin fiebre	
- tiempo de protrombina o tiempo parcial de tromboplastina	1,0	2. Remisión del dolor, las molestias, náuseas y vómitos.	
4. Radiología		B. Con colecistectomía	
a. Rayos X de tórax en el plazo de un año	0,5		
b. Colangiograma intravenoso o colecistograma a menos que el paciente tenga ictericia o vaya a ser operado en 24 horas tras el ingreso	2,0		
B. Colecistitis o colelitiasis crónica			
1. Historia clínica con referencia especial a:			
a. Intolerancia a la comida	1,0		
b. Ataques previos de vesícula	1,0		
c. Ictericia	1,0		

Continúa

Figura 3. Ejemplo de criterios lineales (Continuación)

- | | |
|--|--|
| <p>1. Sin fiebre (99,4 F o menos) a menos que en la historia se justifiquen las causas que justifiquen el alta con temperaturas de 100 F.</p> <p>2. La herida se cura satisfactoriamente</p> <p>3. Recuperación de la función gastrointestinal</p> <p>4. Complicaciones bajo control</p> <p>VII. Período de prehospitalización: con especial atención a:</p> <p>A. Historia clínica: igual que en hospital</p> <p>B. Examen físico: igual que en hospital</p> <p>C. Laboratorio: igual que en hospital</p> <p>D. Radiología: igual que en hospital</p> | <p>VIII. Posthospitalización: con especial atención a:</p> <p>A. Visita al hospital al cabo de una semana o menos y a los seis meses del postoperatorio:</p> <p>1. Examinar la herida</p> <p>2. Registro de la temperatura</p> <p>IX. Finalizar los resultados del tratamiento</p> <p>A. Sin ictericia</p> <p>B. No se repiten los dolores del cuadrante superior derecho o indigestiones</p> <p>C. Sin hernia de incisión</p> |
|--|--|

Fuente: Payne BC, et al. *The quality of Medical Care. Evaluation and improvement*. Chicago: Hospital Research and educational Trustt, 1976. Reimpreso con permiso del editor.

Con el fin de contener costes, la práctica totalidad de los criterios para las estancias hospitalarias especifican el rango de duración aceptable de la estancia, en ausencia de complicaciones; también incluyen una lista de criterios que justifican el ingreso en el hospital, y otra lista que especifica las condiciones en las que el paciente no necesita más atención y debe recibir el alta. Los criterios del alta se basan, claro está, en la descripción de las condiciones clínicas del paciente y en la asunción del rol propio del hospital. Los criterios de ingreso o alta deben ser puramente médicos, aunque en algunos casos se incluyen criterios sociales, tales como la disponibilidad de atención domiciliaria, o la distancia entre el hospital y el domicilio del paciente.

La mayoría de los criterios lineales guardan relación con el proceso de atención

Pocos tienen que ver con los resultados, más allá de los que justifican el alta médica. Es posible, sin embargo, especificar resultados previsibles, al menos en un futuro inmediato. Con ello, se asume que el paciente se compromete con el régimen prescrito y que permanecerá en observación.

Además, si prestamos escasa atención a los resultados de la atención, los criterios lineales no suelen ser lo bastante capaces para especificar contradicciones, y menos capaces todavía de identificar lo innecesario. Se ha planteado que, al especificar únicamente ciertos elementos de atención, los criterios lineales parecen exigir que todo lo que detallan se deba realizar en todos los casos. De este modo, los criterios lineales deben guiar un estado intermedio entre dos extremos no deseados. Si detallan solamente aquello que debe hacerse en cualquier caso, el nivel de calidad de la asistencia que definen, podría resultar muy bajo en muchos casos; si detallan todo aquello que podría necesitarse en un gran número de casos, la asistencia que describen podría resultar innecesariamente costosa en algunos o en muchos casos.

La causa principal de este dilema es la dificultad de ajustar los criterios con sus referentes. Un planteamiento para mejorar el problema, consiste en subclasificar los casos en ca-

tegorías cada vez más exactas y homogéneas, con la correspondiente multiplicación de listas de criterios. Otro enfoque trata de facilitar grandes listados de criterios, especificando un "porcentaje de adhesión" para cada uno de los puntos de la lista, un porcentaje medio o un rango de porcentajes, que indicarían un cumplimiento o adhesión adecuado. Desgraciadamente, este planteamiento sólo puede aplicarse a unos ciertos grupos de casos. Además, es posible, aunque poco probable, estar dentro del rango de porcentajes aprobado gracias a una coincidencia afortunada de errores de omisión y de comisión.

Otro enfoque, aplicable en la evaluación de casos individuales, es introducir "ramificaciones" en los criterios a modo de contingencias. Por ejemplo, no se necesita un test de mononucleosis infecciosa para todos los casos de faringitis, pero sí cuando va acompañado de una inflamación generalizada de los ganglios linfáticos.

Un método ingenioso para incorporar más ramificaciones es subdividir todos los criterios en grupos, uno que incluya elementos de atención o resultados, que deben estar presentes en todos los casos, y otro que incluya a aquellos que no deben estar presentes en ningún caso. Estos criterios de todo o nada van acompañados de listas de excepciones, situaciones en que, en algunos casos individuales (Figura 4)¹ se permite abandonar la dicotomía del todo o nada.

La última etapa de esta progresión es el desarrollo de criterios completamente ramificados o "mapas de criterios". En estos momentos, los criterios ya no parecen listas; adoptan la forma de árboles de decisión o de diagramas de flujo lógico (Figura 5)². El objetivo es la representación aceptable del proceso de manejo de problemas clínicos de manera escalonada, cada paso en función de los resultados del paso previo. Los criterios prescriben una o más acciones aceptables, también subdividen de manera progresiva la cohorte inicial de pacientes en subgrupos cada vez más pequeños, con criterios muy específicos aplicables a todos los miembros de un subgrupo. Así, según van cambiando las propiedades del referente, los criterios cambian al unísono, consiguiendo una adaptación continuada.

Figura 4. Un ejemplo de criterio "Todo o Nada" con excepciones

Criterios para la auditoría de colecistectomía		
Elemento	Estándar 100%	Excepciones
A. Justificación de diagnóstico		
1. Informe positivo de anatomía patológica	100	1a. Ninguna
B. Justificación de Cirugía		
2. Colecistectomía justificada por diagnóstico	100	2a. Dolor grave en el cuadrante superior derecho (CSD) con molestias y espasmos, ictericia de obstrucción o fiebre y leucocitosis; o 2b. Ictericia obstructiva con dolor CSD, molestias y espasmos o fiebre y leucocitosis; o 2c. Historia clínica de dolor CSD, además de rayos X o ecografía positiva por piedras o dos visualizaciones fallidas de vesícula, seguidas de colecistogramas dobles en otro mes; o 2d. "Ileo biliar"; o 2e. Indigestión gaseosa suficiente para interferir en la vida normal, ausencia de otras posibles causas y cálculos biliares demostrado por rayos X o ecografías o dos visualizaciones fallidas de vesícula, seguidas de colecistogramas dobles en otro mes; o 2f. Tumor asintomático de vesícula demostrado en rayos X; y 2g. Segundas opiniones en casos de pacientes con cálculos biliares asintomáticos
3. Exploración de conducto común	0	3a. Historia médica de ictericia o pancreatitis; o 3b. Conducto común dilatado; o 3c. Piedra palpable en el conducto común; o 3d. Pequeñas piedras en vesícula o conducto cístico; o 3e. Coledocograma positivo
4. Cirugía adicional no planificada	0	4a. Apendicetomía incidental
C. Justificación de otros procedimientos		
5. Uso profiláctico de antibióticos	0	5a. Una dosis administrada preoperatoria justo antes de iniciarse la intervención 5b. Se cesa la ingestión de antibióticos pasadas 36 horas de la intervención
6. Transfusión intra o postoperatoria	0	6a. Ninguna
D. Criterios de resultados		
Estado en el alta		
7. La herida está seca y no supura	100	7a. Ninguna
8. Paciente sin fiebre	100	8a. Ninguna
9. Paciente o "allegado" explican la dieta y los niveles de actividad	100	9a. Ninguna
10. Mortalidad	0	10a. Ninguna
E. Indicadores		
Duración de la estancia		
11. Total = 8 días	100	11a. Complicaciones que prolongan la estancia
12. Preoperatorio = X días	100	12a. Ninguna
Complicaciones		
13. Infección de herida quirúrgica	0	13a. Incisión y drenaje y; o 13b. Antibióticos adecuados a I cultivo y sensibilidad
14. Deshiscencia de sutura	0	14a. Ninguna
15. Piedra retenida en conducto común	0	15a. Realización bien de colangiografía intraoperatoria o exploración del conducto común
16. Atelectasis	0	16a. Manejo preventivo 16b. Terapia de inhalación
17. Otras complicaciones	0	17a. Ninguna
Otros indicadores		
18. Ingreso previo para la misma condición o similar	0	18a. Intervención quirúrgica planificada en el primer ingreso
19. Reingreso para cualquier condición similar	0	19a. Ninguna

Fuente: Johson JD. Rationales criteria for a surgical audit of cholecystectomy. Reeditado por *Qual Rev Bull* 1978;4:3-8

Ponderación

La mayoría de los criterios no están ponderados, lo que significa que cualquier punto tiene la misma importancia que cualquier otro. Los criterios también pueden recibir una valoración distinta, en base a enjuiciar aquellos puntos que contribuyen, en mayor o menor medida, en el éxito del manejo. No conocemos de momento la mejor manera de ponderar los criterios, ni tampoco sabemos con certeza, si la ponderación altera las posiciones relativas de los casos o de los proveedores sanitarios en la escala de calidad³.

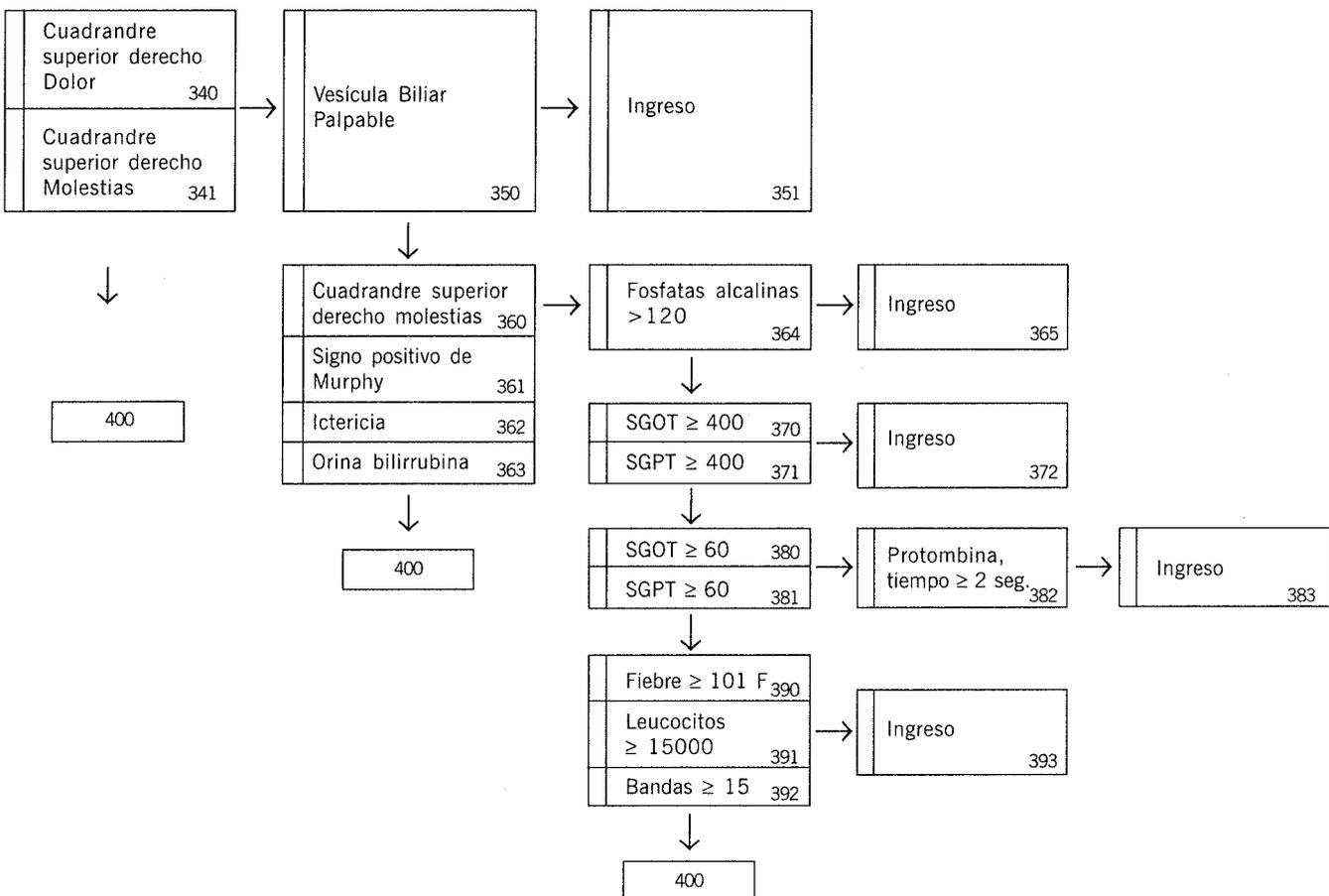
Características de Evaluación

Entre los atributos de evaluación de los criterios, el más importante es la validez (Figura 6). De manera ideal, la validez deriva de documentar de forma seria y científica la relación entre el proceso de asistencia y sus consecuencias. Si esta

relación está bien establecida podemos utilizar como criterios, los elementos del proceso que contribuyan significativamente en la consecución de los resultados deseados. Por la misma razón, podemos utilizar como criterios la consecución de resultados, con la certeza razonable (asumiendo que en ciertas precauciones se establecen el diseño y análisis del estudio) de que los resultados alcanzados son atribuibles a una asistencia previa. Si no se establece una relación entre el proceso y el resultado, no podemos decir que ciertos resultados se deben a la atención dispensada. Simplemente no podemos evaluar la calidad de la atención con ninguna certeza.

Cuando las pruebas científicas en cuanto a la validez no son del todo concluyentes, se opta, a menudo, en su lugar, o para complementar la documentación científica, por un acuerdo entre expertos. Este consenso, además de contribuir a un tipo de validez, confiere una legitimidad profesional a los criterios, haciéndolos más aceptables para clasificar y evaluar la práctica.

Figura 5. Mapa de criterios para el dolor abdominal cuyo origen puede ser la colecistitis*



*Se trata tan sólo de una parte de un mapa más extenso superior, para el dolor abdominal que especifica las indicaciones para el ingreso al hospital. Este extracto se inicia con una nota acerca del dolor o molestias en el cuadrante superior derecho y pretende identificar la presencia de colecistitis, hepatitis, o cualquier otra situación de gravedad en la zona de la vesícula. En el mapa general, la colecistitis se identifica también a través de otros procedimientos. Los criterios están numerados para facilitar el seguimiento del mapa.

Una segunda característica de la evaluación de criterios, es su importancia o relevancia en el manejo de los casos y en los resultados de la atención. Más allá de la confianza o de la preponderancia del juicio de los expertos, para incluir o excluir ciertos criterios de proceso, sabemos poco de la manera de ponderar los criterios y de las consecuencias de esa ponderación. En cuanto a las mejoras en la salud atribuibles a una atención previa, su importancia debe basarse, en gran medida, en las circunstancias y los valores de los pacientes individuales. No podemos profundizar más en esto, pues nos llevaría al difícil terreno de la medición del estado de salud, tema que excede el ámbito de este artículo.

La capacidad de ser documentada es la tercera característica de evaluación de los criterios. Todo el trabajo de evaluación de calidad depende de la disponibilidad y calidad de las historias clínicas. En especial, la omisión de anotar los "resultados negativos" en la historia y la exploración física, ha creado un vacío importante, así como la escasez de información a partir de la cual, se puede deducir la información que recibió el paciente y como se manejó la relación social entablada por médico y paciente. Como resultado de estas dificultades, los criterios de calidad seleccionados para la evaluación están hechos a medida y se limitan a la información de la historia clínica. Llega el momento de considerar las posibilidades de modificar las historias clínicas para poderla adaptar mejor a la evaluación de la calidad.

La adaptabilidad a la variabilidad entre los casos, incluso dentro de un subgrupo de diagnóstico ya se ha descrito como un aspecto del formato. Su reaparición aquí demuestra que la distinción entre las características descriptivas y las características evaluativas es bastante arbitraria. Las características descriptivas se caracterizan únicamente porque afectan a la actuación de los criterios, y las características de evaluación tienen su equivalente en las formas que adoptan los criterios.

Los criterios varían también en cuanto a su estabilidad, es decir, su susceptibilidad de sufrir modificaciones con el tiempo. Por ejemplo, el proceso de elaboración de una historia clínica o de realizar una exploración física experimenta pocos cambios con el tiempo, aunque en otros aspectos de la asistencia no es así. Los criterios de atención deben modificarse adecuadamente. A algunos les preocupa que ciertos criterios, una vez adoptados, se conviertan en inamovibles, evitando la introducción de innovaciones útiles. Puede argumentarse, contrariamente, que si los criterios se revisan constantemente por parte de los mejores médicos, su diseminación podría contribuir a un abandono más rápido de lo obsoleto y a una adopción más temprana de lo nuevo, aunque nunca antes de determinar previamente su valor. De este modo, los criterios pueden emplearse como valiosas herramientas de formación. No obstante, también es cierto, que los criterios se pueden utilizar para impedir la innovación y restringir la práctica. Ciertamente, los que controlan los criterios, tienen también mucha parte del control del sistema asistencial.

Otra característica de la evaluación de criterios es la rigurosidad, o el nivel de calidad que consideran aceptable. Los criterios pueden ser tan estrictos, que la práctica totalidad de casos no cumplen, o tan indulgentes que casi todo se acepte. Existe, obviamente, un aspecto intermedio apropiado para cada

Figura 6. La característica evaluativa más importante de los criterios es la validez

Algunas características de valoración de criterios

- A. Validez
 - 1. Científica
 - 2. Consensuada
- B. Importancia, relevancia
 - 1. Para el manejo del caso
 - 2. Para los resultados
- C. Posibilidad de registros
- D. Adaptación a las variaciones entre los distintos casos
- E. Estabilidad, Oportunidad
- F. Rigor
- G. Eficiencia como elemento de screening

situación. Sin embargo, si el objetivo es la mejora continua, el nivel de calidad exigido debe aumentar gradualmente.

El análisis de la eficiencia es el atributo final de la evaluación. Aparece en escena cuando los criterios explícitos simplificados, tanto del proceso como del resultado, se utilizan como primer paso para separar los casos, entre los que requieren una revisión mayor y más intensivo y los que no. En estas situaciones, la eficiencia del análisis viene determinada por el porcentaje de casos catalogados como cuestionables, que más tarde demuestran no serlo (error de comisión) y el porcentaje de casos aceptados que deberían haber sido cuestionados (error de omisión). Un "screening" muy estricto permite detectar un mayor número de casos cuestionables, aunque también es más probable que detecte casos que, más tarde, se demuestra que se habían tratado adecuadamente. Además de crear un sentimiento de práctica inadecuada, esta estrategia es costosa, ya que se detecta un mayor número de casos que hay que revisar con detalle. Pero si los criterios de "screening" son menos rigurosos, pueden conducir a errores en la dirección contraria, permitiendo que mucha asistencia cuestionable no se detecte y, consecuentemente, se desacredite el programa. Lo ideal, por supuesto, es elaborar los criterios de análisis de modo que se limiten los errores, tanto de comisión como de omisión. Aunque, en el momento de elegir, el grado de rigurosidad dependerá de las penalizaciones asociadas a cada tipo de error: costes, daños morales o lesiones en los pacientes. Desgraciadamente, algunas de estas penalizaciones son muy difíciles de cuantificar.

Procedimientos para Formular

Criterios y Estándares

La Figura 7 muestra los pasos más importantes en los procedimientos habituales de formulación listas de criterios y estándares.

El primer paso, por regla general, es reunir a uno o más grupos de expertos, a cada uno de ellos se le confían unos criterios y estándares para un conjunto de temas. El grado y la naturaleza de los conocimientos son, naturalmente, las primeras consideraciones para seleccionar a los miembros del panel. Es también importante tener en cuenta el modo de representar los distintos segmentos de las profesiones relevantes: por ejemplo, si incluyen a médicos académicos, a médicos de otros centros, medicina general, así como a especialistas y representantes de las principales organizaciones profesionales.

A continuación, hay que seleccionar un método para alcanzar un acuerdo dentro de cada panel. Normalmente, cada panel se reúne en comité tradicional bajo la presidencia de uno

de sus miembros. A veces, no obstante, se utilizan métodos más complejos.

En el método "Delphi", los miembros no suelen reunirse: se les consulta individualmente por correo o telefónicamente. Las opiniones se recogen, se tabulan y se analizan, quizás, en mayor profundidad y los resultados se devuelven a cada miembro. Entonces se les pregunta si se reafirman o cambian su opinión inicial, basándose en el conocimiento de las opiniones de sus colegas. Tras una o más rondas de este tipo, las opiniones tienden a estabilizarse y la encuesta se da por terminada⁴.

El método del "Grupo Nominal" requiere normalmente que los miembros del panel se reúnan. Se les pide que faciliten una opinión por escrito sin consultar al resto. Posteriormente, se tabulan y muestran las opiniones de cada uno de los miembros. A continuación, se abre un debate general sobre los resultados, para que los miembros puedan explicar, elaborar y, quizás justificar sus puntos de vista. Tras el debate general, se pide a cada uno de los miembros que manifieste su opinión final, una vez más, sin consultar al resto⁵.

Figura 7. El proceso de formulación de criterios y estándares se inicia con la selección de un panel de expertos

Procedimientos para formular criterios y estándares

- A. Selección de paneles de expertos
 1. Experiencia
 2. Representatividad, influencia
- B. Selección del método para obtener consenso
 1. Comité tradicional
 2. El método Delphi
 3. El método del "Grupo Nominal"
- C. Selección de referentes
 1. Naturaleza de los referentes
 2. Criterios de selección
 - a. Importancia
 - "Máximo beneficio alcanzable"
 - Administrativa, política, etc.
 - b. Representatividad
 - Muestreo proporcional
 - Muestras ilustrativas (el "método trazador")
 - c. Viabilidad
- D. Selección de estándares de criterios
 1. Conjuntar un grupo que los incluya
 2. Valoración de los apartados
 - a. Importancia
 - b. Posibilidad de registro
 3. Selección de un subconjunto consensuado
 4. Coeficiente de ponderación (si se desea)
 5. Especificación de una ventana temporal
 6. Especificar los estándares alcanzables
- E. Especificar las fuentes de información, los procedimientos y las reglas para la abstracción de información, las variables que deben estandarizarse para poder comparar entre varios casos, etc.
- F. Realización de tests a través de la implementación piloto: viabilidad, fiabilidad, aceptabilidad, eficacia en el "screening" o revisión médica, etc.

Elección de Temas

La primera decisión que deben adoptar los paneles de expertos puede ser la selección de los fenómenos que van a ser objeto de evaluación. Ya se ha discutido el amplio rango de fenómenos que pueden servir como referencia para los criterios y las implicaciones, de elegir una u otra categoría de criterios. Otras consideraciones que pueden influir en la elección de aspectos específicos dentro de una categoría, son los diagnósticos individuales o las condiciones individuales. La importancia de un diagnóstico o condición como objetos de evaluación, viene determinada por la frecuencia con que se producen, por la frecuencia con que suelen manejarse mal y sus consecuencias y las posibilidades de rectificar problemas derivados de un mal manejo. En resumen, la elección se basa en el principio de "máximo beneficio posible atribuible a las mejoras en calidad"⁶. Otros factores, administrativos y políticos pueden tener también a su importancia.

Aparte de la importancia, la elección de diagnósticos, condiciones u otro tema, puede estar guiada por la necesidad de representar razonablemente bien la calidad de la atención en sus distintos segmentos o localizaciones. Podría ser necesaria para algunos propósitos un auténtico muestreo estratificado; es más normal ser más selectivo. El "método de los trazadores"⁷ entra en este tipo de selección. Para implementar este método, hay que construir una matriz que represente los elementos del sistema de atención que se desea evaluar. Por ejemplo, una celda en esta matriz podría representar una atención preventiva para bebés y niños; otra, detección del cáncer en mujeres; y una tercera, los servicios de rehabilitación para hombres de edad más avanzada. Para cada una de estas celdas se seleccionan uno o más diagnósticos o condiciones. En este caso, uno podría elegir inmunizaciones y pruebas de visión en niños, el estadio en que se diagnosticó primero el cáncer de mama en las mujeres, y el manejo y las consecuencias de los ictus en hombres de mayor edad. Cada uno de estos referentes se conoce como "trazador" porque su estudio puede mostrar la actuación del sistema con respecto a alguna función o responsabilidad específicas.

La factibilidad es otro factor en la elección del tema. Los temas elegidos deberían prestarse a la medición casos en que

la ciencia de la atención está avanzada, se dispone de información válida, se pueden formular criterios válidos, los casos se pueden ubicar en categorías, los resultados no dependen demasiado de la colaboración del cliente, etc.

Elección de criterios

Tras seleccionar los temas, puede iniciarse la elección de criterios. Como primer paso, sería útil disponer de una selección amplia donde elegir. La lista inicial puede elaborarse a partir de listas individuales facilitadas por cada uno de los miembros del panel o incluso, por un grupo mayor de informadores y a partir de listas disponibles en bibliografía o que se utilizan en otros lugares.

A continuación, cada aspecto de la lista puede recibir una valoración por parte de los miembros del panel, de manera individual o a través de consulta, utilizando, por ejemplo, una escala del 0 al 100. Estas valoraciones pueden realizarse en función de la importancia para una buena atención o relevancia de unos buenos resultados, o de ambos. Los criterios también se pueden valorar con arreglo a su capacidad para ser encontrados en las historias clínicas. Con el empleo de estas valoraciones, el conjunto principal se puede reducir a un subconjunto de criterios que han recibido valoraciones altas por parte de un gran porcentaje de informadores⁸⁻¹⁰.

Entonces, a los aspectos seleccionados para su inclusión en la lista final se les puede asignar distintos pesos o dejarlos sin ponderar. Luego se asignan los estándares específicos referentes a cada criterio, nuevamente en función de un procedimiento acordado y de reglas para establecer un consenso de grupo.

En especial, cuando se formulan criterios y estándares de resultados, es importante especificar el momento en que se han de evaluar los resultados. El "tiempo ventana" que hay que seleccionar es el que hará que los resultados indiquen las diferencias de calidad de la atención previa y que es factible en vista de las dificultades de obtener información posterior a la conclusión de la asistencia¹¹. No hace falta decir que, para empezar, los criterios y estándares específicos adoptados pueden haber estado influenciados por el tiempo ventana que se espera que la información esté disponible.

Aunque para esta etapa, los criterios y estándares habrán tomado forma, todavía quedará mucho por hacer antes de que puedan implementarse completamente. Se necesita decir, por ejemplo, dónde obtener la información relevante para cada criterio, preparar instrucciones detalladas para abstraer la información de las historias clínicas y para especificar las variables que habrá que tener en cuenta para estandarizar el "case mix". Le sigue un período de pruebas de aplicación a pequeña escala para descubrir la dificultad, el coste, la fiabilidad, la aceptabilidad y eficiencia del método de evaluación. Si se tienen en cuenta las dificultades de implementación reveladas por la prueba piloto, los criterios se podrán ofrecer para un uso más generalizado.

Usos de los Criterios

Embarcarse en un debate sobre los usos de los criterios, podría llevarnos por unos caminos que no pretendemos explorar a través de este artículo. No obstante, puede resultar útil ofre-

cer una perspectiva de lo que hay más adelante, para que los criterios no aparezcan como niños huérfanos sin hogar.

Los criterios y estándares son esenciales para cualquier tipo de evaluación: tanto consciente como inconscientemente, cada actividad de evaluación depende de criterios previos, implícitos o explícitos o ambos. Como la evaluación, es una parte constante de cualquier actividad humana; el empleo de criterios adquiere el correspondiente grado de universalidad y variedad.

Dentro de una perspectiva más estrecha de la calidad, según la entendemos, habitualmente podría diferenciarse entre el uso de criterios para investigación y para la monitorización operativa. La investigación puede adoptar la forma de observaciones epidemiológicas de la distribución de la calidad entre los proveedores sanitarios y los receptores de asistencia¹². Con menor frecuencia, puede adoptar la forma de alteraciones experimentales del modo en que se organiza, financia y suministra la atención, para que podamos descubrir los efectos de estas alteraciones. En ambos casos, tanto si el estudio es epidemiológico o experimental, hay que acordar previamente los fenómenos que indican calidad, y los criterios/estándares que aportan una información más específica acerca de estos indicadores. El objetivo, por supuesto, es adquirir el conocimiento exacto de aquello que nos conduce a la calidad y de aquello que no lo consigue, con el objetivo de poder actuar adecuadamente en la política pública, el diseño de programas y la práctica administrativa.

El uso de criterios en la monitorización de actuaciones (operativa) se puede producir en niveles distintos y en muchos lugares, siendo lo más habitual, la actuación de una institución juzgada según la atención facilitada a sus clientes por parte de los que trabajan en ella. Con referencia a la institución, la evaluación puede ser "externa" o "interna".

La evaluación externa se puede producir en intervalos durante el proceso de acreditación o certificación. Para este propósito, los indicadores de calidad tienden a convertirse en rasgos de la estructura física y de organización de la institución, y los criterios y estándares se formulan de acuerdo con ello. Por ejemplo, si se observa que la manera de organizarse del personal médico de un hospital influye en la calidad, los criterios y estándares deben especificar qué es lo que constituye una forma aceptable de organización.

Aparte de las evaluaciones periódicas con propósitos de acreditación o certificación, la evaluación externa se puede producir de manera más continuada mediante la supervisión a cargo de organizaciones como las agencias estatales de bienestar, aseguradoras u organizaciones de revisión estatales. Este sistema de monitorización se parece, en muchos casos, a la monitorización "interna", que se espera realicen las instituciones a iniciativa propia.

En las operaciones internas de una institución, los criterios interpretan papeles distintos. Representan principalmente, una vez acordados y compartidos, el consenso normativo que imparte objetivos y coherencia a las contribuciones de los distintos participantes de esta labor colectiva. Engloban los objetivos que hacen viable la "gestión por objetivos" tanto en el ámbito administrativo como clínico. Son una herramienta para la formación continuada. Son, claro está, la base de cualquier sistema de monitorización continua.

Los criterios y estándares forman parte del proceso de monitorización, permitiendo, desde el principio la obtención de información detallada sobre si la asistencia ha estado por debajo de los niveles aceptables y, en caso afirmativo, de qué modo. En el punto en que se detectan las deficiencias, es necesario determinar sus causas, para concebir e implantar medidas efectivas y adoptar las medidas necesarias. Aunque las medidas correctivas pueden ser variadas, los esfuerzos de formación centrados de manera específica en las deficiencias detectadas, son la consecuencia más frecuente. Así, los criterios y estándares cumplen la función educativa ya mencionada. Tras implantarse las medidas, los criterios y estándares aparecen en escena como los medios para verificar el éxito o el fracaso de esa medida.

El diseño de un sistema de monitorización, interno o externo, requiere mucho juicio y capacidad. Es necesario, por supuesto, empezar con una visión clara del significado de la calidad y de sus elementos. En el ámbito de la atención técnica, por ejemplo, la calidad implica, entre otras cosas, el diagnóstico preciso y el tratamiento adecuado. Si estos son indicadores de calidad, hay que especificar en detalle, mediante criterios y estándares, lo que requieren e implican un diagnóstico preciso y un tratamiento adecuado.

Existe, no obstante, otra categoría de indicadores de calidad. Se trata de los atributos o sucesos que pueden indicar la probable aparición de deficiencias en el manejo. El "índice de sospecha" aumenta, por ejemplo, cuando la estancia del paciente en el hospital es demasiado corta o demasiado larga, si transcurre demasiado tiempo entre el ingreso y la realización de la intervención prevista, si ciertos procedimientos o intervenciones quirúrgicas se realizan con poca o demasiada frecuencia, o si el paciente vuelve a ingresar en el hospital a poco de recibir el alta. Un sistema eficaz de monitorización servirá para reconocer el posible significado de estos hechos y facilitará los estándares para reconocerlos y los criterios para investigarlos y evaluarlos.

En todas estas actividades de evaluación y de otros muchos modos que deben omitirse por la brevedad de este artículo, los criterios y estándares tienen una presencia primordial. Son, a la vez, un criado dócil y un amo exigente. Ambos mandan y actúan.

Bibliografía

1. Jacobs CM, Christoffel TH, Dixon N. *Mesasuring the Quality of Patient Care: The Rationale for Outcome Audit*. Cambridge, Mass: Ballinger. 1976.
2. Greenfield S, et al. Peer review by criteria mapping: Criteria for diabetes mellitus: The use of decision-making in chart audit. *Ann Intern Med* 1975;83:761-70.
3. Lyons TF, Payne BC. The use of item weights in assessing physician performance with predetermined criteria indices. *Med Care* 1975;13:432-9.
4. Dalkey NC, et al. *The Quality of Life; Delphi Decision-Making*. Lexington. Mass: Lexington Books, 1972.
5. Delbecq AL, Van de Ven AH. A group process model for problem identification and program planning. *Journal of Applied Behavioral Science* 1971;7:466-92.
6. Williamson JW. Formulating priorities for quality assurance activity: Description of a method and its application. *JAMA* 1978;239:631-7.
7. Kessner DM, Kalk CE, Singer J. Assessing health quality. The case for tracers. *N Engl J Med* 1973;288:189.
8. Osborne CE, Thompson HC. Criteria for evaluation of ambulatory child health care by chart audit: Development and testing of a methodology. *Pediatrics* 1975;56(suppl):625-92.
9. Riedel RL, Riedel DC. *Practice and performance: An Assessment of Ambulatory Care*. Ann Arbor. Mich: Health Administration Press, 1979.
10. Hulka BS, et al. Peer review in ambulatory care: Use of explicit criteria and implicit judgments. *Med Care* 1979;17(suppl):1-73.
11. Brook RH et al. Assessing the quality of medical care using outcome measures: An overview of the method. *Med Care* 1977;15(suppl):1-165.
12. Donabedian A. The epidemiology of quality. *Inquiry* 1985;522:282-302.
13. Lembcke PA. Medical auditing by scientific methods. Illustrated by major female pelvic surgery. *N Engl J Med* 1956;162:645-55.