

Calidad de vida. Algunas consideraciones metodológicas

J. A. Fernández-López, R. Hernández-Mejía
y A. Cueto-Espinar

Área de Medicina Preventiva y Salud Pública.
Departamento de Medicina. Facultad de Medicina.
Universidad de Oviedo. Oviedo (Asturias).

El desarrollo de sistemas para medir los efectos económicos de los trastornos de salud en los trabajadores durante la segunda mitad de siglo XX ha ido evolucionando conceptualmente, desde el estado de salud después del tratamiento, en los años 40 (¿puede el paciente seguir trabajando?), a la afectación de la calidad de vida en los 80 (¿puede afectar el tratamiento a la calidad de vida del paciente?).

Las medidas de la salud por sí mismas no tienen implicaciones en la actividad médica; sin embargo, como proceso en sí, permite que se alcancen ciertos objetivos que previamente eran inobtenibles. Esto se deriva de dos importantes aspectos que caracterizan al proceso: primero, permiten medir el impacto de una determinada enfermedad en la comunidad, y segundo, y más importante, proporcionan el medio por el cual puede medirse el impacto de las intervenciones terapéuticas, médicas o sociales sobre la salud anterior de los individuos, evaluando así la efectividad de las intervenciones.

El término «medición» en ciencias médicas, a menudo implica la idea de una operación precisa, basada en procedimientos bien establecidos y expresada en términos de medida estandarizada. Este escenario contrasta fuertemente con los intentos de las ciencias sociales para desarrollar medidas de salud donde no sólo el fenómeno bajo estudio está definido de diferentes modos, sino que varían las opciones de cómo deberían representarse y, por supuesto, cuantificarse. La consecuencia ha sido un largo e intenso esfuerzo, en las últimas décadas, para desarrollar estas medidas en su construcción y metodología.

Teoría de la medición de la calidad de vida

De la propia indefinición de calidad de vida (CDV) se desprende la dificultad de su cuantificación. Fanshel y Bush sugirieron la posibilidad de determinar una tabla de puntuación para cada día, que representara la calidad de vida de un individuo. La suma de estas puntuaciones

a lo largo del tiempo daría lugar al llamado Índice Acumulado de Salud (HSI). De esta manera describieron once estados, en términos de función/disfunción, que resultan clásicos en la literatura médica, y que abarcan desde el estadio 1 (bienestar completo) hasta el 11 (muerto).

Con fines prácticos podría ser conveniente referirse a CDV en términos de una simple escala conceptual y deducir de ello un índice simple combinando cantidad con calidad de vida. Un indicador así sería el instrumento donde diferentes tratamientos son comparados o donde se apoya la toma de decisiones en la clínica práctica. De hecho, el análisis coste-efectividad en la investigación de los servicios y economía de la salud ha desarrollado para este caso el indicador AVAC (años de vida ganados ajustados por calidad) —Quality-Adjusted Life-Years (QALY)— donde la duración de la vida ganada por un específico tratamiento médico es ajustada por pesos de «utilidad», los cuales reflejan el valor relativo de un estado de salud comparado con otro.

Recientemente ha habido una tendencia creciente entre los profesionales de la salud, investigadores y economistas a reconocer la necesidad de desarrollar métodos para medir y cuantificar el cambio en el estado de salud como resultado de una actividad médica dada. En el logro de este objetivo se han desarrollado tres enfoques principales: a) primero, implica el uso de escalas numéricas *ad hoc*; b) el segundo, considera la adecuación entre lo que se paga y lo que se recibe; c) el tercero, a través del uso de utilidades y AVAC o QALY.

El uso de escalas numéricas *ad hoc* implica medir al individuo en una serie de aspectos de su salud, asignando una puntuación numérica a cada medición y sumando los puntos. Karnofsky en 1949 utilizó esta aproximación en su índice y posteriormente lo hicieron Grogono y Woodgate en 1971, en su Índice de Medición de Salud, y Spitzer en el suyo en 1981.

Mishan desarrolló el enfoque de «pagar de buena gana» (WTP) (*willingness to pay*), que está firmemente basado en la moderna economía del bienestar. Se basa en la idea de que las valoraciones de los individuos se refle-

jan en lo que ellos están dispuestos a pagar, para recibir ciertos beneficios o evitar ciertos costes (pagar es usado aquí en el sentido de que los individuos están dispuestos a anteponer o sacrificar y no solo en el sentido monetario).

El tercer enfoque para medir el estado de salud, iniciado por Torrance, es a través del uso de «utilidades» y AVAC. Se apoya en el empleo de una escala cardinal en la que las diferencias entre los valores individuales a lo largo de la escala pueden ser comparados de una manera significativa. Un ejemplo diario de tal escala es el uso de grados centígrados para las medidas de temperatura. Del mismo modo, Torrance describe «utilidades» como los valores cardinales que se corresponden a cada estado de salud, en una escala en la que se asigna un valor 1,0 (completamente sano) y 0,0 (estar muerto). El centro de atención radica en la determinación de los pesos numéricos o valores de utilidad.

Ante esta perspectiva cabe preguntarse: ¿son las utilidades válidas? Estos valores deberían ser científicamente medibles mediante instrumentos fiables y válidos que puedan ser usados con el público general. Con esto en mente, Torrance realizó una investigación empírica de tres de las técnicas de medida más comúnmente usadas: método de categorías (*Rating Scales*), que aplica la técnica de grados; la técnica de riesgo estándar (*Standard Gamble*) y la técnica del compromiso temporal (*Time Trade-off*); esta última resultó ser la mejor de las tres. De este modo, por ejemplo, se juzgó que un año con efectos secundarios derivados de un tratamiento antihipertensivo es equivalente a 0,98 de año sano.

De manera resumida, se puede decir que la valoración de la calidad de vida comprende las funciones físicas, emocionales y sociales del individuo y pasa por un proceso de información, esquematizado muy didácticamente por Patrick y Erickson, que comienza por la definición del problema, sigue con la conceptualización del estatus de salud, la selección del instrumento, la planificación y realización del trabajo de campo, el análisis y la presentación de resultados.

Requerimientos metodológicos de las mediciones

En la mayoría de los casos el sistema de medida ha sido especificado por los investigadores mismos, conforme a un arbitrario pesaje de su propio criterio (medidas *ad hoc*). Al igual que cualquier método científico de medición, los instrumentos de calidad de vida han de reunir ciertos criterios metodológicos:

La *factibilidad* de cada técnica fue determinada por tres criterios principales: su aceptación por los jueces, su facilidad de uso por los entrevistadores y su coste.

La *comparabilidad* de las tres técnicas fue medida en términos de si producían los mismos valores o no, y si

no, si los valores derivados están relacionados de algún modo sistemático que permite la construcción de curvas de conversión.

En cuanto a la *validez/fiabilidad* de las mediciones, si las técnicas de medida realmente miden lo que se pretende medir, en este caso la utilidad o fuerza de las preferencias de un individuo por ciertos estados de salud, se dice que son válidas. En sus *Standards of Educational and Psychological Tests*, la *American Psychological Association*, define tres tipos básicos de validez: de criterio, de contenido y de constructo o estructura.

La validez de criterio se alcanza si la medida propuesta corresponde a la más exacta medida observable de un fenómeno. Evidentemente tales estándares raramente existen en el campo de la construcción de calidad de vida. La validez de un instrumento puede ser ampliamente definida como la habilidad para medir lo que se quiere medir. Esto es difícil de lograr en áreas como aislamiento social o satisfacción de la vida, puesto que no existen modelos estandarizados de comparación. La evidencia de validez es usualmente obtenida examinando las relaciones de las puntuaciones finales de una variedad de medidas.

La validez de contenido se refiere a la facultad de los ítems del instrumento para representar adecuadamente el área que pretenden medir. La selección del contenido es guiada por una construcción teórica escogida de un índice de calidad de vida. Sin embargo, es difícil definir los requerimientos metodológicos apropiados para evaluar el contenido de la validez. Por ello, el mayor tipo de validación, usado en las ciencias sociales y de la conducta en general y en estudios de CDV en particular, es la validez de constructo.

La validez de *constructo* o estructura considera tres métodos diferentes: los tests estadísticos de las dimensiones incluidas en la construcción (análisis de factor); la covariación de estas dimensiones respecto a medidas independientes del mismo fenómeno (validez convergente) y la variación esperada de las mediciones de CDV respecto a criterios como edad, sexo y estado de salud (validez discriminatoria).

Una excelente revisión de los índices y sus características ha sido realizado por McDowell y Newell. Ellos ilustran desacuerdos metodológicos sofisticados entre los investigadores de este campo, pero demuestran que existe un fuerte cuerpo de evidencia respecto de la validez convergente y discriminatoria. Considerando conjuntamente los requerimientos de las medidas en cuanto a contenido, convergencia y validez discriminatoria, los datos pueden constituir estándares empíricos de la investigación actual sobre calidad de vida.

La aplicación de los resultados de calidad de vida, en la práctica, debe considerar diferencias entre los pacientes según edad y sexo. En concordancia con resultados previos, se asumió que las mujeres tenían resultados infe-

riores que los hombres en dimensiones como la vitalidad, y más altos en ansiedad y depresión, después de ajustadas las variaciones de la edad. Se asume que la edad ejerce un importante efecto independiente en todos los resultados; los análisis de varianza y covarianza mostraron cambios significativos en la dirección esperada. La repetibilidad/fiabilidad es una medida de la consistencia de la información y se establece normalmente por resultados de tests-retests. Puede ser expresada a través de una variedad de métodos estadísticos como correlación de coeficientes, valor kappa o desviación estándar de mediciones repetidas o de diferencias de mediciones idénticas. En ensayos realizados, las puntuaciones finales de calidad de vida pueden resultar de una elevada confianza cuando una situación es constante, en orden a ser expresión de los cambios en las mismas debidos a los tratamientos. El grado de variación entre diferentes observadores puede también ser medido en ensayos en los cuales es utilizado más de un entrevistador; evidentemente este error se minimiza en el caso de utilizar un solo entrevistador durante todo el ensayo.

El aspecto relativo a la *sensibilidad* de las medidas ha sido menos desarrollado. En medicina clínica, detectar mejoría o empeoramiento del estado de salud, asociado a la enfermedad y especialmente con el tratamiento aplicado, resulta de crucial importancia. La sensibilidad, en este contexto, es la habilidad de un método para detectar cambios a lo largo del tiempo, en concordancia con las condiciones de enfermedad o tratamiento. Las medidas han de ser suficientemente fidedignas y válidas para permitir conclusiones acerca del cambio. Al mismo tiempo, las medidas han de ser tan finas que permitan diferenciar distintos niveles de bienestar, por ejemplo de un tratamiento respecto de otro. La sensibilidad precisa ser mostrada mediante comparación de cambios en las puntuaciones entre grupo control y grupo de tratamiento.

Todos los instrumentos de medición deben estar diseñados para detectar cambios «menores», en el rango «normal» del bienestar humano. A mayor sensibilidad del test corresponde un mayor «ruido psicológico». Estos cambios pueden ser debidos a variaciones espontáneas o endógenas en el bienestar o debidas a efectos secundarios a tratamientos o de la medida utilizada. A pesar de todo esto, en general puede afirmarse que el aspecto de la sensibilidad precisa más consideración en la futura metodología sobre calidad de vida.

La información sobre CDV siempre está referida, en alguna medida, a una valoración subjetiva de los sufrimientos de la enfermedad o su tratamiento. Las características de la enfermedad y su tratamiento son importantes determinantes de la variabilidad y selectividad de las dimensiones de la CDV. Esto se ilustra con los siguientes ejemplos: en un estudio con pacientes de enfermedad terminal cardíaca, se observó alteración en

las relaciones familiares en la mitad, mientras que el 90% de ellos tenían actividades y hábitos sociales normales. Esto nada tiene que ver con la situación que viven los afectos de hipertensión arterial leve o límite. No es adecuado describir estas diferencias exclusivamente como cambios a lo largo de la misma dirección. Ante todo, deben introducirse diferentes mediciones para describir la CDV de acuerdo con el tipo de enfermedad. Otro ejemplo se refiere al cáncer de mama. Es conocido que no existe diferencia en cuanto a la supervivencia, entre cirugía total o parcial de la mama; sin embargo, ha sido detectada una ganancia en las medidas de CDV para el caso de una cirugía más conservadora. Las dimensiones más relevantes de la CDV en pacientes con cáncer de mama son la imagen corporal y la depresión. De nuevo habría que decir que no es factible examinar los efectos de la imagen corporal en pacientes sin impacto físico aparente.

Las características esenciales de las mediciones psicométricas en investigaciones de CDV (confianza, sensibilidad, validez, fiabilidad, etc) han sido discutidas en distinta forma por diversos autores.

Métodos de medición de calidad de vida

Los métodos de medida de salud han sido llamados medidas medicosociales y cubren una variedad de aspectos que abarcan desde la incapacidad física, el dolor y el bienestar emocional y social, a la satisfacción y la calidad de vida. Algunos autores consideran que, tanto clínicos como investigadores, desconocen frecuentemente la amplia gama de técnicas de medida disponibles para su uso. No se usan frecuentemente las mejores técnicas disponibles ni se aplican métodos comparables, lo que explica la poca evidencia científica acumulada. En general, el desarrollo de las mediciones en este campo fue lento y realizado de forma aislada por quienes deseaban hacer uso de ellas.

El usuario de tales mediciones debe decidir, en primer lugar, lo que quiere medir; debe ver si se trata de evaluar un programa de cuidados o estudiar individualmente a los sujetos, si el fenómeno en estudio es agudo o crónico y cómo de amplia ha de ser la medida y de detallada la información.

A lo largo de 100 años la mayoría de los países occidentales han recogido información estadística, caracterizando condiciones sociales. Los datos describiendo tasas de nacimiento y muerte, educación, criminalidad, vivienda, empleo y condiciones económicas reflejaron problemas de preocupación pública y, a menudo, han sido punto de partida de movimientos de reforma. Las medidas de salud han formado siempre un componente central de esa preocupación pública, siendo usadas para afrontar los mayores problemas de la sociedad, contribuir al proceso de alcanzar éxitos políticos y a monito-

rizar la efectividad de medidas de intervención médica y de salud.

La evolución de las sociedades y la alteración de los problemas de salud han creado la necesidad de nuevos indicadores que reflejen los resultados de ese cambio. En cada caso la resolución de un problema de salud puso de manifiesto otros problemas que reducen la utilidad del indicador existente y, así, el círculo comienza de nuevo.

Los cuestionarios constituyen la base de la mayoría de las mediciones de la calidad de vida y pretenden reflejar la evaluación de una enfermedad y/o su tratamiento por medio de los sufrimientos que se originan. En la construcción de un cuestionario de CDV juega un papel esencial la selección de los ítems que han de constituir las diferentes escalas de bienestar. Cada uno de estos modelos desarrollados sigue las siguientes etapas: selección de ítems, reducción del número de ítems, formato del cuestionario, pretestificado, validez, repetibilidad y sensibilidad.

Existe un amplio rango de métodos para recoger información base acerca del paciente, tales como entrevista, consulta ordinaria, entrevistas con personal al caso, entrevistas telefónicas y tests «lápiz-papel». Una medición aproximada, a partir de un estudio específico de las respuestas categorizadas, ha llevado al uso de escalas análogas visuales, especialmente en áreas como el dolor y el humor. Las escalas visuales análogas asumen que la intensidad y la frecuencia de un síntoma pueden ser representadas sobre una línea continua y que esta representación es entendida por el paciente de la misma manera.

Otra medición aproximada la constituyen las escalas tipo Likert; en este caso se asume que cada ítem puede ser contestado en un punto «n» de la escala (usualmente del 1 al 5), donde el punto final define el resultado de un indicador y donde cada número representa un aspecto diferente del fenómeno bajo estudio, mediante transcritores como «nunca» y «todo el tiempo». Como Guyatt apuntó: «una escala Likert es bien entendida por los diferentes tipos de pacientes según su educación y requiere un proceso instructivo mínimo».

En la metodología actual se tienen en cuenta la construcción de la escala y el análisis de su consistencia interna, así como la aplicabilidad, la validez y confianza y la sensibilidad al cambio. Como se ha apuntado anteriormente, la selección del instrumento de medida viene determinada por la enfermedad bajo estudio, por el diseño y las hipótesis del mismo, por los requerimientos psicométricos y por las consideraciones prácticas de tiempo y dinero.

Tipos de cuestionario

Los cuestionarios, que son la base de la mayoría de los estudios de CDV, pueden ser administrados por un en-

trevistador o ser cumplimentable por el propio encuestado. Ambos tipos tienen ventajas e inconvenientes, que se pueden resumir como sigue:

Cuestionarios cumplimentables por el mismo encuestado

Esta elección tiene como primera desventaja el que excluye pacientes que no pueden leer o escribir, bien por razones educacionales, culturales o de salud, especialmente en gente mayor. La cumplimentación del instrumento, para que sea completa, precisa que sus instrucciones sean claras; asimismo deben estar diseñados para diferentes niveles de instrucción e inteligencia. Existen fórmulas disponibles que previenen puntuaciones para diferentes lecturas y habilidades, basadas en el contenido del texto.

Otra desventaja de este método es que cuando no están completamente cumplimentados no es posible saber si se debe a una omisión por defectos de comprensión o si el sujeto no desea dar esa información. Estos cuestionarios son ampliamente usados en estudios clínicos amplios multicéntricos, para los cuales el entrenamiento, la estandarización y el despliegue de un equipo entrevistador no es factible. Como ventajas tiene el ser reproducible y barato, así como su posible utilización a través del correo. Algunos de los más conocidos se muestran en la tabla 1.

Cuestionarios administrados con entrevistador

La principal ventaja del uso del entrevistador es la menor restricción existente entre quien pregunta y lo que puede ser preguntado. Esto resulta de importancia capital en las mediciones de CDV, donde es muy poco conocido el impacto de la enfermedad y/o su tratamiento en la vida de los pacientes.

La apertura a las preguntas permite al enfermo la oportunidad de describir sus percepciones de problemas, sus expectativas y sus mejoras. Es conocido también que el entrevistador puede usar maneras específicas de obtener respuestas detalladas. La calidad de los datos obtenidos por entrevista depende del entrevistador; así, la habilidad principal de un entrevistador es establecer un

TABLA 1
Algunos ejemplos de medidas de salud general autoadministrables

The Sickness Impact Profile (SIP)
The McMaster Health Index Questionnaire
The Psychological General Well-Being (PGWB) Index
The General Health Rating Index
The General Health Policy Model/Quality of Well Being Scale (QWB)
The Nottingham Health Profile (NHP)

buen reportaje con los entrevistados, sin influir en sus respuestas y asegurándose que el entrevistado conteste todas las preguntas correctamente.

Los entrevistadores médicos pueden tener problemas para adoptar un papel neutral en la entrevista, especialmente cuando se trata de evaluar resultados de tratamientos. Cuando son precisos equipos de entrevistadores, la variabilidad entre ellos debe ser reducida mediante sesiones de entrenamiento y debe calcularse la repetibilidad. Ciertos sesgos pueden estar fuera del control de los investigadores: se ha mostrado que la raza, la edad y el sexo de los investigadores influye en las respuestas.

Cuando se va a estudiar la CDV en personas de edad, debemos emplear un entrevistador. Los ancianos pueden ser inexpertos en cumplimentar cuestionarios, tener una visión pobre o una escritura dificultosa. Si el entrevistador se encuentra un paciente sordo puede utilizar una lista de fichas de respuestas para su elección.

Los cuestionarios disponibles para medir CDV se encuentran en su mayoría en lengua inglesa, y salvo el *Nottingham Health Profile* (NHP) —del cual existe una versión en español y catalán— y el *Sickness Impact Profile* (SIP), que tiene una versión en español, no es posible aplicarlos directamente en nuestro país. El asunto de los métodos de CDV en diferentes países y su aplicación transcultural es de gran preocupación. En los primeros estudios se vieron diferencias en cuanto al bienestar de salud percibida entre nacionalidades, principalmente inmigrantes en EE. UU de descendentes irlandeses, comparados con italianos. Posteriormente se han usado varios instrumentos de CDV en diferentes países occidentales mostrando, en general, resultados similares.

Dado que los cuestionarios no han sido desarrollados suficientemente en nuestro país, deben ser adaptados, desde otras lenguas, mediante traducciones y contraducciones a un lenguaje aceptable para la persona de la calle. En general se logran aproximaciones muy exactas, pero excepcionalmente puede no ser posible encontrar un significado equivalente de algún aspecto concreto entre dos lenguas diferentes.

Según la áreas que cubra el cuestionario de la CDV, se definirán sus usos y aplicaciones. En general no puede

aplicarse un cuestionario si no se trata del campo para el cual ha sido diseñado. Por ejemplo, perfiles como el SIP o el NHP, que describen limitaciones en el estilo de vida producidas por la enfermedad, tales como pérdida de autocuidado o movilidad, no son adecuados para ser utilizados en pacientes con hipertensión, los cuales están libres de síntomas antes del diagnóstico y, a lo sumo, pueden experimentar leves efectos psicológicos del etiquetado. Sin embargo, cuando la hipertensión se complica en forma de ataque o insuficiencia cardíaca con la consiguiente pérdida de independencia, estos cuestionarios pueden ser útiles.

No existe todavía una medida válida estándar en el campo de la CDV, a pesar de la múltiple variedad de medidas desarrolladas, pero el hecho de que la naturaleza multidimensional biopsicosocial de la CDV haya encontrado acuerdo entre los investigadores y el desarrollo de las diferentes mediciones físicas, emocionales y sociales llevadas a cabo en las últimas décadas nos permiten hacer el operacional el concepto y cada día está más cerca la posibilidad de obtener un instrumento válido y cómodo para ser usado en la clínica diaria logrando una mejor evaluación terapéutica de los pacientes y desde la epidemiología obtener un más refinado conocimiento de la salud de las poblaciones.

Bibliografía recomendada

- Alonso Caballero J, Anto Boque JM. Instrumentos de medida de la calidad de vida relacionada con la salud: características y proceso de adaptación transcultural. *Quadern CAPS* 1990; 14:16-24.
- Elinson J, Siegman AE. *Socio-medical health indicators*. New York: Baywood Publishing Company, Inc, 1979.
- Fletcher AE, Bulpitt CJ. Measurement of quality of life in clinical trials of therapy. *Cardiology* 1988; 75 (suppl 1):41-52.
- McDowell I, Newell C. *Measuring Health: a guide to rating scales and questionnaires*. New York: Oxford University Press, 1987.
- Siegrist J, Junge A. Background material for the workshop on QALYs. Conceptual and methodological problems in research on the quality of life in clinical medicine. *Soc Sci Med* 1989; 29:463-468.
- Spilker B. *Quality of life in assessment in clinical trials*. New York: Raven Press Ltd, 1990.
- Teeling Smith G. *Measuring Health: a practical approach*. London: John Wiley & Sons, 1988.
- Van Dam F. Quality of life: methodological aspects. *Bull Cancer* (París) 1986; 73:607-613.
- Walker SR, Rosser RM. *Quality of life: assessment and application*. Lancaster: MTP Press L, 1988.
- Wenger NK, Mattson ME, Fuberg CD. *Assessment of quality of life in cardiovascular therapies*. New York: Le Jacq Publishing, 1984.