

El Cuestionario de Salud SF-36 español: una década de experiencia y nuevos desarrollos

Gemma Vilagut^a / Montse Ferrer^a / Luis Rajmil^b / Pablo Rebollo^c / Gaietà Permanyer-Miraldà^d / José M. Quintana^e / Rosalía Santed^a / José M. Valderas^a / Aida Ribera^d / Antonia Domingo-Salvany^a / Jordi Alonso^{a,f},
por los investigadores de la Red-IRYSS*

^aUnidad de Investigación en Servicios Sanitarios. Institut Municipal d'Investigació Mèdica (IMIM-IMAS). Barcelona. España.

^bAgència d'Avaluació de Tecnologies i Recerca Mèdica (AATRM) de Catalunya. Barcelona. España.

^cUnidad de Investigación en Resultados de Salud. Servicio de Nefrología. Hospital Universitario de Asturias. Oviedo. España.

^dUnidad de Epidemiología. Servicio de Cardiología. Hospital Universitario Vall d'Hebron. Barcelona. España.

^eUnidad de Investigación. Hospital de Galdakao. Vizcaya. España.

^fUniversitat Autònoma de Barcelona (UAB). Barcelona. España.

(The Spanish version of the Short Form 36 Health Survey: a decade of experience and new developments)

Resumen

Objetivo: El Cuestionario SF-36 es uno de los instrumentos de Calidad de Vida Relacionada con la Salud (CVRS) más utilizados y evaluados. Tras una década de uso este artículo revisa críticamente el contenido, propiedades métricas y nuevos desarrollos de la versión española.

Métodos: Revisión de los artículos indexados en Medline (PubMed) y en las bases de datos IBECs e IME que han utilizado la versión española del cuestionario. Se seleccionaron los artículos con información sobre modelo de medida, fiabilidad, validez y sensibilidad al cambio del instrumento.

Resultados: Se encontraron 79 artículos, 17 de los cuales describían características métricas del cuestionario. En el 96% las escalas superaron el estándar propuesto de fiabilidad (α de Cronbach) de 0,7. Las estimaciones agrupadas obtenidas por metaanálisis fueron superiores a 0,7 en todos los casos. El SF-36 mostró buena discriminación entre grupos de gravedad, correlación moderada con indicadores clínicos y alta con otros instrumentos de CVRS. El SF-36 predijo mortalidad y detectó mejoría tras la angioplastia coronaria, la cirugía de hipertrofia prostática benigna o la ventilación domiciliar no invasiva. Los nuevos desarrollos descritos (puntuaciones basadas en normas, la versión 2, el SF-12 y el SF-8) mejoraron sus propiedades métricas y su interpretación.

Conclusiones: El SF-36, conjuntamente con las nuevas versiones desarrolladas, es un instrumento muy adecuado para su uso en investigación y en la práctica clínica.

Palabras clave: SF-36. Calidad de vida relacionada con la salud. Validez. Fiabilidad. Cuestionarios.

Abstract

Objective: The Short Form-36 Health Survey (SF-36) is one of the most widely used and evaluated generic health-related quality of life (HRQL) questionnaires. After almost a decade of use in Spain, the present article critically reviews the content and metric properties of the Spanish version, as well as its new developments.

Methods: A review of indexed articles that used the Spanish version of the SF-36 was performed in Medline (PubMed), the Spanish bibliographic databases IBECs and IME. Articles that provided information on the measurement model, reliability, validity, and responsiveness to change of the instrument were selected.

Results: Seventy-nine articles were found, of which 17 evaluated the metric characteristics of the questionnaire. The reliability of the SF-36 scales was higher than the suggested standard (Cronbach's alpha) of 0.7 in 96% of the evaluations. Grouped evaluations obtained by meta-analysis were higher than 0.7 in all cases. The SF-36 showed good discrimination among severity groups, moderate correlations with clinical indicators, and high correlations with other HRQL instruments. Moreover, questionnaire scores predicted mortality and were able to detect improvement due to therapeutic interventions such as coronary angioplasty, benign prostatic hyperplasia surgery, and non-invasive positive pressure home ventilation. The new developments (norm-based scoring, version 2, the SF-12 and SF-8) improved both the metric properties and interpretation of the questionnaire.

Conclusions: The Spanish version of the SF-36 and its recently developed versions is a suitable instrument for use in medical research, as well as in clinical practice.

Key words: SF-36. Health-related quality of life. Validity. Reliability. Questionnaires.

Correspondencia: Jordi Alonso. IMIM-IMAS.
Doctor Aiguader, 80. 08003 Barcelona. España.
Correo electrónico: jalonso@imim.es

*Red de investigación cooperativa para la Investigación en Resultados de Salud y Servicios Sanitarios.

Recibido: 31 de marzo de 2004. *Aceptado:* 22 de noviembre de 2004.

Introducción

El cuestionario de salud SF-36 fue desarrollado a principios de los noventa, en Estados Unidos, para su uso en el Estudio de los Resultados Médicos (Medical Outcomes Study, MOS)¹. Es una escala genérica que proporciona un perfil del estado de salud y es aplicable tanto a los pacientes como a la población general. Ha resultado útil para evaluar la calidad de vida relacionada con la salud (CVRS) en la población general y en subgrupos específicos, comparar la carga de muy diversas enfermedades, detectar los beneficios en la salud producidos por un amplio rango de tratamientos diferentes y valorar el estado de salud de pacientes individuales². Sus buenas propiedades psicométricas, que han sido evaluadas en más de 400 artículos³, y la multitud de estudios ya realizados, que permiten la comparación de resultados, lo convierten en uno de los instrumentos con mayor potencial en el campo de la CVRS.

Algunas de estas razones impulsaron a realizar la adaptación para su uso en España. Desde la primera publicación de este proceso de adaptación⁴, el SF-36 y su versión reducida de 12 ítems, el SF-12, se han convertido en un instrumento muy útil en la evaluación de resultados en nuestro medio. Tras casi una década de experiencia de muchos investigadores españoles es necesario hacer un balance sobre las potencialidades y las realidades del cuestionario. En este artículo se revisan el contenido del cuestionario, el proceso de su adaptación al español y las evidencias sobre sus características métricas. También se hace una valoración sobre las aplicaciones y las limitaciones detectadas, así como un resumen de los desarrollos novedosos del cuestionario que intentan superarlas.

Contenido del SF-36

El Cuestionario de Salud SF-36 está compuesto por 36 preguntas (ítems) que valoran los estados tanto positivos como negativos de la salud. Se desarrolló a partir de una extensa batería de cuestionarios empleados en el MOS, que incluían 40 conceptos relacionados con la salud². Para crear el cuestionario, se seleccionó el mínimo número de conceptos necesarios para mantener la validez y las características operativas del test inicial. El cuestionario final cubre 8 escalas, que representan los conceptos de salud empleados con más frecuencia en los principales cuestionarios de salud, así como los aspectos más relacionados con la enfermedad y el tratamiento².

Los 36 ítems del instrumento cubren las siguientes escalas: Función física, Rol físico, Dolor corporal, Salud general, Vitalidad, Función social, Rol emocional y Salud

mental. Adicionalmente, el SF-36 incluye un ítem de transición que pregunta sobre el cambio en el estado de salud general respecto al año anterior. Este ítem no se utiliza para el cálculo de ninguna de las escalas pero proporciona información útil sobre el cambio percibido en el estado de salud durante el año previo a la administración del SF-36⁵.

Hay 2 versiones del cuestionario en cuanto al período recordatorio: la «estándar» (4 semanas) y la «aguda» (1 semana). El cuestionario está dirigido a personas de ≥ 14 años de edad y preferentemente debe ser autoadministrado, aunque también es aceptable la administración mediante entrevista personal y telefónica⁵. La consistencia interna no presentó diferencias entre los cuestionarios autoadministrados y los administrados mediante entrevista⁵.

Desarrollo de las versiones españolas del SF-36

La traducción al español del cuestionario ha sido descrita con detalle⁴. Se siguió un protocolo común en los países participantes en el proyecto internacional de adaptación del cuestionario original, el International Quality of Life Assessment (IQOLA)^{6,7}. El cuestionario original desarrollado en Estados Unidos fue traducido al español por 2 personas bilingües cuya lengua materna era el español. Cada uno de ellos realizó una traducción independiente de los ítems del cuestionario y de las opciones de respuesta. Además, puntuaron la dificultad de traducción en una escala de 0 (ninguna dificultad) a 100 (extremadamente difícil). Posteriormente, los traductores se reunieron con el investigador principal (J.A.) para acordar una traducción común, debatir las diferencias y documentar las posibles alternativas y las decisiones tomadas. La traducción consensuada se entregó a 2 evaluadores que puntuaron su calidad en una escala de 0 (totalmente inadecuada) a 100 (perfecta), según 3 criterios: claridad, utilización de lenguaje común y equivalencia conceptual. A partir de sus valoraciones se consensuó una traducción directa preliminar que se entregó a 2 nuevos traductores, cuya lengua materna era el inglés norteamericano. Éstos obtuvieron una nueva versión en inglés (traducción inversa), que fue comparada con la versión original para evaluar la equivalencia conceptual. Posteriormente, se realizó una reunión con los autores de las versiones ya disponibles en otros países en la que se trató de armonizar el contenido del cuestionario. Finalmente, se realizaron estudios piloto con diversos grupos de pacientes crónicos para valorar la comprensión y la factibilidad de administración del cuestionario.

Antes de producir la versión definitiva del cuestionario, se llevó a cabo un estudio empírico de calibración de sus opciones de respuesta, que demostró la ordinalidad de éstas, esto es, la validez de su ordena-

ción o secuencia, y una gran equivalencia con la versión original americana^{4,7}.

Los investigadores del proyecto IQOLA se esforzaron por intentar mantener la comparabilidad de las diferentes versiones al adaptar culturalmente los ítems. Así, en el caso de actividades físicas regulares concretas, seleccionaron actividades culturalmente apropiadas que capturaran de forma adecuada el gasto global de energía involucrado en la actividad y, al mismo tiempo, representarían la utilización de las mismas partes del cuerpo y pudieran llevarse a cabo indistintamente por mujeres y varones. Por ejemplo, la actividad de jugar al golf (ítem PF02) de la versión original se tradujo por ir en bicicleta en Italia y Holanda, caminar por el bosque o por jardines en Suiza, y caminar durante más de 1 h en España⁸.

Asimismo, hay una versión en catalán que fue desarrollada a partir de la española, mediante una doble traducción y consenso entre los traductores y el equipo investigador. También hay una versión en eusquera, obtenida siguiendo un método muy similar al descrito para la versión española, y que ha mostrado su factibilidad y validez⁹.

Puntuación de las escalas

Las escalas del SF-36 están ordenadas de forma que a mayor puntuación mejor es el estado de salud. La tabla 1 contiene el número de ítems incluidos en las diferentes escalas del cuestionario, así como una breve descripción del significado de puntuaciones

altas y bajas⁵. Para el cálculo de las puntuaciones, después de la administración del cuestionario, hay que realizar los siguientes pasos:

1. Homogeneización de la dirección de las respuestas mediante la recodificación de los 10 ítems que lo requieren, con el fin de que todos los ítems sigan el gradiente de «a mayor puntuación, mejor estado de salud».

2. Cálculo del sumatorio de los ítems que componen la escala (puntuación cruda de la escala).

3. Transformación lineal de las puntuaciones crudas para obtener puntuaciones en una escala entre 0 y 100 (puntuaciones transformadas de la escala).

Así pues, para cada dimensión, los ítems son codificados, agregados y transformados en una escala que tiene un recorrido desde 0 (el peor estado de salud para esa dimensión) hasta 100 (el mejor estado de salud) (tabla 1). En caso de que falte información, si se han contestado al menos el 50% de los ítems de una escala, los autores recomiendan sustituir cualquier ítem ausente por el promedio de los ítems completados de ésta. En caso contrario (más del 50% de ítems no contestados), la puntuación de dicha escala no se debería calcular. Además el cuestionario permite el cálculo de 2 puntuaciones sumario, la componente sumario física (PCS) y la mental (MCS), mediante la combinación de las puntuaciones de cada dimensión (tal como se detalla a continuación).

Los estudios de análisis factorial realizados en diferentes muestras americanas^{5,10,11}, y sobre muestras representativas de la población general de 9 países eu-

Tabla1. Contenido de las escalas del SF-36

Dimensión	N.º de ítems	Significado de las puntuaciones de 0 a 100	
		«Peor» puntuación (0)	«Mejor» puntuación (100)
Función física	10	Muy limitado para llevar a cabo todas las actividades físicas, incluido bañarse o ducharse, debido a la salud	Lleva a cabo todo tipo de actividades físicas incluidas las más vigorosas sin ninguna limitación debido a la salud
Rol físico	4	Problemas con el trabajo u otras actividades diarias debido a la salud física	Ningún problema con el trabajo u otras actividades diarias debido a la salud física
Dolor corporal	2	Dolor muy intenso y extremadamente limitante	Ningún dolor ni limitaciones debidas a él
Salud general	5	Evalúa como mala la propia salud y cree posible que empeore	Evalúa la propia salud como excelente
Vitalidad	4	Se siente cansado y exhausto todo el tiempo	Se siente muy dinámico y lleno de energía todo el tiempo
Función social	2	Interferencia extrema y muy frecuente con las actividades sociales normales, debido a problemas físicos o emocionales	Lleva a cabo actividades sociales normales sin ninguna interferencia debido a problemas físicos o emocionales
Rol emocional	3	Problemas con el trabajo y otras actividades diarias debido a problemas emocionales	Ningún problema con el trabajo y otras actividades diarias debido a problemas emocionales
Salud mental	5	Sentimiento de angustia y depresión durante todo el tiempo	Sentimiento de felicidad, tranquilidad y calma durante todo el tiempo
Ítem de Transición de salud	1	Cree que su salud es mucho peor ahora que hace 1 año	Cree que su salud general es mucho mejor ahora que hace 1 año

ropeos del proyecto IQOLA^{12,13}, demostraron que las 8 escalas definen 2 componentes principales de salud, la componente sumario física (PCS) y la mental (MCS). Los pasos principales de los algoritmos de cálculo de las puntuaciones de las 2 medidas sumario son: a) estandarización de las 8 escalas del SF-36 con la media y la desviación estándar (DE) de la población general; b) ponderación de las escalas aplicando los pesos factoriales obtenidos en un análisis de componentes principales, y c) agregación de las escalas y transformación para obtener una media de 50 y una DE de 10 en la población general.

Las puntuaciones de 0 a 100 de las escalas del SF-36 han sido ampliamente utilizadas y gozan de popularidad por la traducción directa de sus máximo y mínimo a el mejor y el peor de los estados de salud posibles. Sin embargo, los autores proponen para las nuevas componentes sumario del SF-36 las puntuaciones basadas en normas, cuya principal ventaja es que los resultados son directamente interpretables respecto a la población de referencia¹⁴. Así, puntuaciones superiores o inferiores a 50 indican mejor o peor estado de salud, respectivamente, que la media de la población de referencia.

Los investigadores del proyecto IQOLA⁶ evaluaron y compararon 2 maneras de calcular las puntuaciones sumario en otros países: a) algoritmos de cálculo específicos para cada país, o b) algoritmos de cálculo estándar, derivados de la población general americana¹⁵. Se observó una concordancia muy elevada entre los 2 tipos de cálculo de las puntuaciones (correlaciones superiores a 0,98), por lo que se recomienda, en estudios nacionales, el uso de algoritmos de cálculo específicos que facilite la interpretación respecto a la población general del propio país y de los algoritmos de cálculo estándar en estudios internacionales, para permitir la comparación directa entre países. En ese caso, las puntuaciones sumario deberán interpretarse en relación con los valores obtenidos en la población general de Estados Unidos.

Métodos

Se realizó una revisión de la bibliografía nacional e internacional sobre la versión española del SF-36. Para ello, se llevó a cabo una búsqueda bibliográfica en la base de datos Medline (PubMed) de los términos conjuntos «SF-36» o «SF36», y «Spanish» o «Spain», incluyendo también este último como término del tesoro por estar explícito en el árbol temático de localización geográfica. También se realizó una búsqueda de los términos «SF-36» o «SF36» o «calidad de vida» en todos los índices de la base bibliográfica española IBECS, y de «SF-36» o «SF36» en la base de datos del IME (ti-

tulo, descriptores y texto libre). La búsqueda se llevó a cabo el 14 de julio de 2004.

Propiedades métricas evaluadas

Siguiendo los criterios de evaluación de instrumentos de CVRS del Medical Outcome Trust¹⁶, se recogieron los datos sobre la fiabilidad (consistencia interna y reproducibilidad), el modelo de medida, la validez de constructo y la sensibilidad al cambio. Fueron recogidos los coeficientes α de Cronbach y de correlación intraclass (CCI), que evalúan la consistencia interna y la reproducibilidad test-retest, respectivamente. Para el α de Cronbach, se ha establecido el valor 0,7 como punto de corte mínimo para comparaciones de grupos, y el de 0,9 para comparaciones individuales¹⁷. El CCI tiene un rango de 0 (ninguna concordancia) a 1 (concordancia perfecta), y un valor superior a 0,75 se considera como acuerdo excelente¹⁸. Para estudiar el modelo de medida se recogieron los resultados de análisis factoriales. La validez de constructo fue evaluada mediante comparaciones de grupos previamente definidos con diferencias esperables en CVRS, y correlaciones del SF-36 con indicadores clínicos y con otros cuestionarios de CVRS. La convergencia interna y discriminante de los ítems se evalúan mediante las correlaciones de cada uno de los ítems con la dimensión a la que teóricamente representan y con el resto de las escalas, respectivamente. El porcentaje de éxito de escalamiento es una medida resumen que consiste en la proporción de ítems dentro de una escala que presentan correlaciones más elevadas con su propia escala que con las del resto del cuestionario. El tamaño del efecto (TE) es un estimador de la magnitud del cambio y se calcula como la diferencia entre las medias (antes y después de la intervención) dividida por la DE antes de la intervención^{19,20}:

$$TE = (m1 - m2)/s1$$

El tamaño del efecto traduce los cambios a una unidad de medida estándar que permite la comparación entre diferentes medidas: 0,20 (pequeño), 0,5 (moderado) y 0,80 (grande)^{20,21}.

Metaanálisis

Se obtuvo una estimación conjunta del coeficiente de consistencia interna a partir de un metaanálisis de los coeficientes α de Cronbach de los estudios evaluados. Se evaluó la heterogeneidad entre los diferentes estudios mediante el test de la χ^2 , fijando un riesgo α de error de tipo I de 0,10 para determinar la significación del estadístico (en lugar del habitual valor de 0,05 debido al bajo poder de esta prue-

ba cuando el número de estudios es pequeño)²². Según el resultado del test de heterogeneidad, se reportan los estimadores sumarios del metaanálisis a partir del modelo de efectos fijos de Mantel-Haenszel o del modelo de efectos aleatorios DerSimonian and Laird²³.

Resultados

Se identificaron 93 artículos, de los que fueron excluidos 13 por ser realizados con otras versiones en castellano (principalmente versiones hispanas de Estados Unidos), y uno por haber utilizado una versión modificada *ad hoc*²⁴. De los 79 artículos restantes, mediante evaluación del título y el sumario, se identificaron 11 que evaluaban alguna de las características métricas del cuestionario^{4,12,13,25-32}. Dado que no se detectó ningún estudio longitudinal específicamente diseñado para evaluar las características métricas, se incluyeron también los 6 estudios longitudinales detectados³³⁻³⁸ que valoraban la capacidad predictiva de mortalidad del SF-36, y/o tratamientos de eficacia demostrada para estimar la sensibilidad al cambio. En total, se recogió información de 17 artículos, realizados sobre 15 estudios diferentes, de los cuales 7 estaban publicados en revistas españolas. La tabla 2 contiene una breve descripción de los diferentes estudios seleccionados, que fueron llevados a cabo tanto en la población general (5 artículos) como en pacientes con diferentes trastornos (12 artículos). Los artículos están ordenados según el ámbito poblacional. De los 15 estudios evaluados, 9 de ellos eran transversales, 4 eran longitudinales con evaluación de la calidad de vida antes y después de una intervención terapéutica y los 2 restantes eran estudios observacionales de seguimiento. En 7 de los artículos se proporcionaron estadísticos descriptivos sobre la distribución de las escalas, 9 proporcionaron información sobre la fiabilidad, 9 evaluaron la validez del cuestionario y, adicionalmente, 4 de los artículos seleccionados permitían evaluar la sensibilidad al cambio (tabla 2).

Fiabilidad

En la tabla 3 se detallan los resultados de fiabilidad del cuestionario evaluada a partir de los coeficientes α de Cronbach (consistencia interna) y los CCI (reproducibilidad).

En la mayoría de los estudios, el coeficiente de consistencia interna α de Cronbach superó el valor mínimo recomendado para las comparaciones de grupos (α de Cronbach = 0,7) en todas las escalas, excepto en la Función social. Las escalas Rol físico, Función física y Rol emocional obtuvieron los mejores resulta-

dos de fiabilidad y en la mayoría de ocasiones superaron el valor de 0,90, límite recomendado para las comparaciones individuales.

Con el fin de obtener una estimación conjunta del coeficiente α de Cronbach de cada escala a partir de todos los estudios, se llevó a cabo un metaanálisis para cada una de las escalas utilizando siempre el modelo de efectos aleatorios²³, dado que se observó heterogeneidad entre los estudios en todos los casos. Las estimaciones conjuntas obtenidas a partir del metaanálisis de los coeficientes α de Cronbach fueron $\geq 0,9$ para las escalas Función física, Rol físico y Rol emocional. Las estimaciones conjuntas del resto de las escalas superaron el valor de 0,7.

La reproducibilidad del cuestionario se evaluó mediante la comparación de las puntuaciones medias obtenidas en las 2 administraciones del cuestionario (t de Student) y el cálculo de los CCI en el estudio realizado en pacientes con enfermedad coronaria estable. Las puntuaciones medias de la administración inicial y la realizada a las 2 semanas fueron muy similares, con una diferencia marginalmente significativa para la dimensión de Función física (-3,7; $p = 0,04$). Los CCI fluctuaron entre 0,58 en la escala de Rol emocional y 0,99 en la escala de Rol físico⁴ (tabla 3).

Éxito de escalaje y efectos suelo y techo

En la tabla 4 se muestran las características de la distribución de las puntuaciones (correlación ítem-dimensión, éxito de escalaje [%], efecto suelo [%] y efecto techo [%]). La consistencia interna de los ítems evalúa si éstos tienen correlaciones elevadas con la escala a la que teóricamente representan. Se recomiendan valores de las correlaciones superiores a 0,4³⁹. Las correlaciones de los ítems con su propia escala fueron superiores al límite mínimo recomendado en los 2 estudios que evaluaron la consistencia interna de los ítems^{25,29}. El éxito de escalaje fue del 100% para todas las escalas en los 2 estudios que evaluaron esta propiedad^{25,29}, excepto para la escala de Función social del estudio de pacientes con enfermedad pulmonar obstructiva crónica (EPOC) (56%).

El efecto suelo del cuestionario fue inferior al 15% para todas las escalas en la mayoría de los estudios, excepto en las escalas de Rol físico y Rol emocional de la muestra de pacientes con EPOC²⁵ y la de pacientes con una posible cardiopatía isquémica²⁹, donde fueron superiores al 17,5%. El valor máximo observado fue del 33,0%. El efecto techo fue inferior al 15% para las escalas de Función física, Salud general, Vitalidad y Salud mental en la mayoría de estudios. Sin embargo, se mantuvo elevado en el resto de las escalas, en especial para las escalas Rol físico y Rol emocional. La versión 2 del SF-36, que se describe al final del capítulo, intenta so-

Tabla 2. Características de los estudios evaluados que incluyen el SF-36 español

Referencia	Muestra	n	Tipo de estudio	Mujeres (%)	Edades	Media de edad	Información descriptiva ^a	Fiabilidad	Validez ^b	Sensibilidad al cambio
Población general										
Alonso et al, 1998 ²⁶	Población general española	9.151	Transversal	51,8%	≥ 18	45,2	X	X	X	
Ware et al, 1998 ¹³	Población general de Santander	1.250	Transversal	50,2%	18-64	ND	X	X	X	
Keller et al, 1998 ¹²	Población general española de ≥ 60 años	3.949	Transversal	56,4%	≥ 60	ND	X	X	X	
Ayuso-Mateos et al, 1999 ²⁷										
López-García et al, 2003 ³²	Pacientes con enfermedad coronaria estable (comprobada por el cardiólogo)	46	Transversal (2 evaluaciones: durante visita y 2 semanas después)	ND	-	ND	X	X	X	
Muestras de pacientes										
Alonso et al, 1995 ⁴	Pacientes con posible cardiopatía isquémica	185	Transversal	28,1%	-	60,2	X	X	X	
Failde et al, 2000 ²⁹	Pacientes después de una intervención de derivación aortocoronaria	710	Longitudinal	14,6%	< 81	63	X	X	X	
Permanyer et al, 2001 ³⁷										
Permanyer et al, 2001 ³⁶	Pacientes tratados con angioplastia coronaria con balón o <i>stent</i>	397	Longitudinal	28,6%	-	63	X	X	X	
Alonso et al, 1998 ²⁵	Pacientes con EPOC, varones	321	Transversal	0%	-	64,9	X	X	X	
Domingo-Salvany et al, 2002 ³³	Seguimiento de pacientes con EPOC	312	Seguimiento	0%	-	65	X	X	X	
Espinosa de los Monteros et al, 2002 ²⁸	Pacientes con asma	219	Transversal	65,2%	> 14	46	X	X	X	
García Ordóñez et al, 2001 ³⁰	Pacientes con VIH en todas las fases de la enfermedad	30	Transversal	23,7%	-	32,9	X	X	X	
Salinas Sánchez et al, 2002 ³⁸	Pacientes con HBP en lista de espera de cirugía	181	Longitudinal	0%	-	68,8	X	X	X	
Nauffal et al, 2002 ³⁵	Pacientes con cifoescoliosis o enfermedad neuromuscular iniciando tratamiento de ventilación domiciliaria no invasiva	62	Longitudinal	43,5%	-	51,1	X	X	X	
Gómez-Besteiro et al, 2004 ³¹	Pacientes en lista de espera de trasplante de riñón o trasplantados en 1997	285	Transversal	36,5	-	47	X	X	X	
López Revuelta, 2004 ³⁴	Pacientes en diálisis renal crónica (julio de 1996 hasta octubre de 1998)	318	Cohorte							
	Evaluaciones: SF-36 basal; estado vital y número de días de hospitalización cada 3 meses			39,6%	-	60,2	X	X	X	

ND: No disponible; EPOC: enfermedad pulmonar obstructiva crónica; VIH: virus de la inmunodeficiencia humana; HBP: hiperplasia benigna de próstata.

^aDatos descriptivos de las escalas. ^bValidez de constructo o validez predictiva.

Tabla 3. Fiabilidad de las escalas del SF-36 español

Autor	Muestra	Escala							
		Función física	Rol físico	Dolor corporal	Salud general	Vitalidad	Función social	Rol emocional	Salud mental
Alonso et al., 1998 ²⁶	Población general española (n = 9.151)	0,94	0,96	0,92	0,82	0,78	0,81	0,94	0,79
Ayuso-Mateos et al., 1999 ²⁷	Población general de Santander (n = 1.250)	0,90	0,86	0,70	0,81	0,71	0,73	0,85	0,79
López-García et al., 2003 ²⁸	Población general española de ≥ 60 años (n = 3.949)	0,93	0,95	0,87	0,79	0,85	0,84	0,91	0,85
Alonso et al., 1995 ⁴	Pacientes con enfermedad coronaria estable (n = 50)	0,84	0,90	0,88	0,71	0,74	0,45	0,94	0,77
	Coefficiente de correlación intraclass	0,83	0,99	0,65	0,85	0,70	0,65	0,58	0,76
Faílde et al., 2000 ²⁹	Pacientes con posible cardiopatía isquémica (n = 185)	0,92	0,91	0,82	0,74	0,78	0,72	0,94	0,80
Alonso et al., 1998 ²⁵	Pacientes con EPOC, varones (n = 312)	0,92	0,90	0,84	0,77	0,86	0,55	0,91	0,85
Espinosa de los Monteros et al., 2002 ²⁸	Pacientes con asma (n = 219)	0,79	0,79	0,80	0,79	0,77	0,78	0,80	0,78
García Ordóñez et al., 2001 ³⁰	Pacientes con VIH (n = 300)	0,91	0,90	0,92	0,77	0,84	0,71	0,87	0,80
Gómez-Besteiro et al., 2004 ³¹	Pacientes en lista de espera de trasplante de riñón o trasplantados en 1997 (n = 285)	0,88	0,92	0,87	0,70	0,82	0,78	0,89	0,82
Estimación conjunta	(IC del 95%) ^a	0,9	0,91	0,86	0,78	0,8	0,74	0,9	0,81
		(0,87-0,92)	(0,87-0,94)	(0,8-0,9)	(0,76-0,8)	(0,76-0,83)	(0,69-0,79)	(0,87-0,93)	(0,78-0,83)

EPOC: enfermedad pulmonar obstructiva crónica; VIH: virus de la inmunodeficiencia humana; IC: intervalo de confianza.

^aObtenida a partir de metaanálisis. Modelo de efectos aleatorios de DerSimonian y Laird²⁵.

Tabla 4. Características de la distribución de las puntuaciones y validez de constructo del SF-36 español

Autor	Muestra	Correlación ítem-dimensión (éxito del escalaje, %) ^a	Efecto suelo (%) ^b	Efecto techo (%) ^b	Modelo de medida	Discriminación entre grupos	Correlación con indicadores
Ware et al, 1998 ¹³	Población general española (n = 9.151)	NR	0,2-13,3	2,8-86,0	Análisis factorial	NR	NR
Keller et al, 1998 ¹²	Población general española (n = 9.151)	NR	NR	NR	Ecuaciones estructurales	NR	NR
Ayuso-Mateos et al, 1999 ⁷	Muestra aleatoria población de Santander (n = 1.250)	NR	0,1-5,7	3,2-89,1	Análisis factorial	Enfermedades crónicas Edad y sexo	GHQ-12
López-García et al, 2003 ²⁰	Población general española de ≥ 60 años (n = 3.949)	NR ^c	0,5-21,5	0,9-78,6	NR	NR	NR
Fáilde et al, 2000 ²⁹	Pacientes con posible cardiopatía isquémica (n = 185)	0,53-0,95 (100%)	0,0-33,0	3,0-65,0	Análisis factorial	Enfermedades crónicas Factores de riesgo Gravedad Edad y sexo	GHQ-28
Alonso et al, 1998 ²⁵	Pacientes con EPOC, varones (n = 312)	0,39-0,85 (56,3-100%)	0,6-20,1	0,0-69,1	NR	Gravedad clínica Edad y sexo	SGRQ Disnea VEMS VEMS PEF Años de evolución
Espinosa de los Monteros et al, 2002 ²⁸	Pacientes con asma (n = 219)	NR	0,0-5,8	0,4-19,8	NR	Gravedad Sexo	NR
Gómez-Besteiro et al, 2004 ³¹	Pacientes en lista de espera de trasplante de riñón o trasplantados en 1997 (n = 285)	NR	0,12-15,8	0,4-77,6	Análisis factorial	NR	NR

^aRango por ítems del cuestionario. ^bRango por dimensión. NR: no reportado; EPOC: enfermedad pulmonar obstructiva crónica; GHQ: General Health Questionnaire; SGRQ: St George's Respiratory Questionnaire; VEMS: volumen espiratorio máximo por segundo.

lucionar este problema de las 2 escalas de rol aumentando el número de categorías de las opciones de respuesta de sus ítems⁴⁰.

Modelo de medida y validez de contenido

El análisis factorial de componentes principales sobre la muestra de 9.151 individuos de la población general¹³ cumplió las hipótesis previas de 2 factores, físico y mental. Tal y como se esperaba, las escalas de Función física, Rol físico y Dolor corporal tuvieron correlaciones más elevadas con la componente física ($r \geq 0,74$), y las escalas Salud mental, Rol emocional y Función social se correlacionaron más con la componente mental ($r \geq 0,62$). Las escalas Salud general y Vitalidad mostraron correlaciones significativas con ambas componentes principales. Este patrón factorial demostró ser estable en diferentes países y en submuestras según grupos de edad y sexo¹³. El análisis de ecuaciones estructurales realizado sobre la misma muestra¹² confirmó las relaciones previamente hipotetizadas entre los ítems del SF-36 y las escalas a las que pertenecen. Los resultados también sugerían un buen ajuste para el modelo considerando dos factores de orden 2 que podían ser interpretados como Salud física y mental, si bien se obtenía un mejor ajuste considerando un factor de orden 2 adicional que podía ser considerado como Bienestar general.

En 3 de los estudios evaluados^{27,29,31} se llevaron a cabo análisis factoriales descriptivos de los ítems del cuestionario. Al menos 5 de las 8 escalas propuestas por los autores del cuestionario original mostraron una correspondencia muy precisa entre los factores y las escalas en cada uno de los estudios, aunque estas 5 escalas no fueron siempre las mismas en todos los análisis.

Validez de constructo

En el estudio sobre 1.250 individuos de la ciudad de Santander²⁷, los individuos de la población general con alguna enfermedad crónica de larga evolución y los que habían consultado recientemente a un médico tenían puntuaciones más bajas, de forma significativa, en todas las escalas del cuestionario. Como se esperaba, se observó una peor percepción del estado de salud en las mujeres para todas las escalas, con diferencias estadísticamente significativas para las escalas de Función física, Rol emocional, Salud mental, Vitalidad y Dolor corporal. Las escalas de Vitalidad, Salud general, Dolor corporal y Función física mostraron correlaciones negativas significativas con la edad. Se observó también una correlación negativa estadísticamente significativa entre el cuestionario de salud mental General Health Questionnaire (GHQ) y las 8 escalas del SF-36 (tabla 4).

En la muestra de 185 pacientes con una posible cardiopatía isquémica²⁹ se confirmaron las hipótesis previas, de tal manera que las puntuaciones de las escalas Función física y Salud general empeoraban conforme aumentaba el número de enfermedades adicionales (comorbilidad) o el número de factores de riesgo, y las mujeres mostraron peores resultados en todas las escalas excepto Dolor corporal. Además, los pacientes con una puntuación ≥ 6 en la versión de 28 ítems del GHQ obtuvieron también peores resultados en todas las escalas del SF-36.

En el estudio de 321 pacientes con EPOC²⁵, la asociación entre las puntuaciones del SF-36 y las medidas de función respiratoria siguieron el patrón de jerarquía esperado. Las correlaciones más altas se obtuvieron entre el SF-36 y el cuestionario específico sobre el impacto de los problemas respiratorios St. George's Respiratory Questionnaire (SGRQ) (0,30-0,81). El valor absoluto de correlación de las escalas del SF-36 con la disnea fue intermedio (rango, 0,19-0,50), superior al rango de correlaciones observado con el VEMS (0-0,44), si bien éstas fueron significativas para 6 de las 8 escalas. En este mismo estudio, las puntuaciones medias de las escalas físicas mostraron un gradiente claro y diferencias significativas según el nivel de gravedad clínica.

Se evaluó también la validez del SF-36 en la muestra de pacientes con asma²⁸ y se observaron diferencias significativas respecto a la calidad de vida según la gravedad del asma en todas las escalas excepto Rol emocional. Se observaron correlaciones significativas con el VEMS para todas las escalas, excepto Rol físico y Rol emocional.

Cabe destacar que las correlaciones de las puntuaciones del cuestionario con las variables fisiológicas y clínicas fueron moderadas o bajas, y siempre mayores con las medidas clínicas (p. ej., la disnea, la capacidad funcional observada o la prueba de esfuerzo) que con los indicadores de fisiopatología (p. ej., el VEMS). Esta evidencia muestra que el SF-36 y los cuestionarios de CVRS, en general, miden un concepto diferente de los que evalúan estos indicadores.

Un estudio reciente evalúa el SF-36 sobre 691 pacientes con distintos diagnósticos médicos⁴¹ (cáncer de pulmón, cáncer hematológico, osteoartritis de cadera, trastornos de la alimentación y enfermedad intestinal crónica). En este estudio se observa la capacidad del cuestionario de discriminar entre las distintas áreas de la salud afectadas en cada uno de los grupos diagnósticos. Así, por ejemplo, en las escalas de Función física, Rol físico y Dolor corporal, las puntuaciones más bajas se presentaban en los grupos de osteoartritis de cadera, cáncer de pulmón y cáncer hematológico. El grupo de individuos con trastornos alimentarios presentó la peor puntuación en la escala de Salud mental, Rol emocional y Función social.

Validez predictiva

El SF-36 se ha mostrado como un predictor independiente de la mortalidad en varios estudios clínicos. El primero se basa en la cohorte de los 321 pacientes con EPOC descritos anteriormente³³. Tras un seguimiento medio de 4,8 años, el sumario físico del SF-36 era predictor de la mortalidad total y respiratoria de los 312 pacientes que pudieron seguirse (*hazard ratio* estandarizada, 1,36 y 1,40, respectivamente). Esta asociación se obtuvo en un modelo de Cox en el que también estaban incluidos la edad, el VEMS y el índice de masa corporal (IMC).

En un estudio en 318 pacientes que iniciaron diálisis renal crónica³⁴, la MCS del SF-36 predijo la mortalidad general tras más de 2 años de seguimiento, una vez ajustados los resultados por todos los indicadores clínicos de gravedad. Además, entre el subgrupo de pacientes diabéticos (n = 208), tanto la MCS como la PCS iniciales predijeron la mortalidad y las hospitalizaciones en ese período.

En la cohorte de 710 pacientes que requirieron cirugía coronaria³⁷, la PCS preoperatoria fue predictora de mortalidad total al año de la intervención, así como de acontecimientos clínicos en los supervivientes iniciales, ajustando por características de riesgo quirúrgico, del proceso asistencial y del acto quirúrgico.

Sensibilidad a los cambios

La sensibilidad a los cambios es una evidencia longitudinal de la validez de los instrumentos de medida (tabla 5). Para el SF-36 español, esta característica ha sido valorada en menor grado que las propiedades mencionadas previamente. El cuestionario de salud SF-36 ha mostrado su capacidad de detectar cambios clínicos en una muestra de 710 individuos que requirieron cirugía coronaria³⁷, donde globalmente los pacientes evaluados experimentaron una mejoría significativa a los 6 meses y al año de la intervención en todas las escalas del SF-36. Al año, los tamaños del efecto fueron de 0,88 para la PCS y de 0,35 para la MCS. Asimismo, en una muestra de pacientes tratados con angioplastia coronaria³⁶, las puntuaciones de todas las escalas habían aumentado significativamente al año de la intervención (tamaño del efecto entre 0,45 y 1,29), con una mejora especialmente notable en las escalas de Rol físico y Dolor corporal (magnitud del efecto de 1,29 y 1,11, respectivamente) (tabla 5).

En la muestra de pacientes con hiperplasia prostática benigna que requirieron cirugía³⁸, las puntuaciones obtenidas 1 año después de la operación mostraron una mejora significativa en todas las escalas del cuestionario. Las magnitudes del efecto fueron mayo-

res de 0,5, lo que denota un cambio moderado, para la escala de Rol físico y para la PCS. En los pacientes con cifoscoliosis que habían sido tratados con ventilación no invasiva domiciliaria³⁵, se detectaron tamaños del efecto superiores a 0,5 (cambio moderado), en la mayoría de las escalas excepto Dolor corporal y Salud general, y superiores a 0,8 (cambio elevado) para las escalas de Rol físico y Rol emocional, a los 3, 6, 9 y 18 meses.

Discusión

Los estudios publicados sobre las características métricas de la versión española del SF-36 aportan suficiente evidencia sobre su fiabilidad, validez y sensibilidad. Estos resultados son consistentes con el instrumento original, lo que demuestra que la versión española es equivalente y puede ser utilizada en estudios nacionales e internacionales.

Por otra parte, es importante destacar que se trata del cuestionario genérico para la medida de la CVRS que goza de mayor vitalidad. Desde su aparición en el inicio de la década de los noventa no ha cesado el diseño de nuevas herramientas de ayuda a la interpretación (valores de referencia poblacionales, valores sumario y diferencia mínima clínicamente relevante, entre otras), así como de nuevas versiones para solucionar algunos de los problemas detectados.

Normas poblacionales de referencia del SF-36 español

En España, los valores de referencia fueron obtenidos a partir de las respuestas al instrumento de 9.151 españoles mayores de 18 años²⁶. El estudio proporciona información sobre la distribución de las puntuaciones de las escalas del SF-36 según el sexo, así como los percentiles de las puntuaciones según los grupos de edad (18-24, 25-34, 35-44, 45-54, 55-64, 65-74, ≥ 75 años) y sexo. En la actualidad se ha ampliado este estudio en la población española de ≥ 60 años de edad³², de manera que se dispone de valores de referencia válidos y precisos hasta edades avanzadas (≥ 85 años). Éste es un hecho especialmente relevante si tenemos en cuenta el envejecimiento poblacional experimentado en nuestra sociedad y la creciente necesidad de información relacionada con la salud en estos grupos de edad.

Hay valores de referencia para las puntuaciones de cada dimensión del SF-36 para la versión original obtenidos a partir de las respuestas a la National Survey of Functional Health Status (NSFHS) de 1990, que incluía el SF-36. Se entrevistaron 2.474 individuos de la población general de ≥ 18 años de Estados Unidos.

Tabla 5. Sensibilidad al cambio. Tamaño del efecto observado en las puntuaciones por escalas y en cada uno de los componentes sumario de los estudios que evaluaban cambio en la calidad de vida entre antes y después de la aplicación de un tratamiento

Autores, año de publicación	Tratamiento aplicado (n.º de pacientes)	Dimensiones	PCS	MCS
Permanyer Miralda et al, 2001 ³⁷	Cirugía coronaria (710)	NR	0,88	0,35
Permanyer Miralda et al, 2001 ³⁶	Angioplastia coronaria (397)	0,45-1,29	NR	NR
Salinas Sánchez et al, 2002 ³⁸	Cirugía de hipertrofia prostática benigna (181)	0,2-0,58	0,5	0,19
Nauffal et al, 2002 ³⁵	Ventilación domiciliar no invasiva (62)	0,11-1,24	NR	NR

NR: no reportado; PCS: componente sumario física; MCS: componente sumario mental.

También hay valores de referencia basados en la población general para otros países, como Alemania, Dinamarca, Francia, Países Bajos, Italia, Japón, Noruega, Reino Unido y Suecia. Además, a partir de las muestras obtenidas en estos países, se han publicado valores para pacientes con trastornos crónicos: alergias, artrosis, diabetes, enfermedad cardíaca isquémica, hipertensión, insuficiencia cardíaca y patología respiratoria crónica⁴².

Nuevos desarrollos del cuestionario de salud SF-36

Versión 2 del cuestionario SF-36

En 1996 se desarrolló la versión 2.0 del cuestionario original SF-36⁴⁰, con el principal objetivo de mejorar las características métricas de las 2 escalas Rol emocional y Rol físico. Las mejoras en la versión 2.0 incluyeron: *a)* instrucciones y enunciados más sencillos para algunas de las preguntas; *b)* mejoras en la pre-

sentación de los ítems y las opciones de respuesta en la versión autoadministrada del cuestionario para facilitar su lectura y cumplimentación, y reducir así el número de respuestas faltantes; *c)* mayor comparabilidad en las diferentes traducciones y adaptaciones culturales del cuestionario; *d)* 5 opciones de respuesta en vez de opciones de respuesta dicotómicas en los ítems correspondientes a las escalas Rol físico y Rol emocional, y *e)* eliminación de una de las 6 opciones de respuesta («Muchas veces») para los ítems de Salud mental y Vitalidad. De forma similar a las medidas sumario PCS y MCS, las puntuaciones de la versión 2 utilizan algoritmos de puntuación basados en las normas poblacionales para las 8 escalas del SF-36 (media ± DE, 50 ± 10 para la población general). El programa de cálculo de las puntuaciones de la versión 2.0 también mejora la estimación de respuestas faltantes².

Actualmente ya se ha desarrollado la versión 2.0 de la familia de cuestionarios SF-36 en español, aunque por el momento no hay datos sobre su validación y, sobre todo, no hay estudios comparativos con la versión 1. La información sobre las características de las 2 ver-

Tabla 6. Características de las diferentes versiones del cuestionario

Puntuaciones	N.º de ítems			N.º de niveles				Puntuación*	
	SF-36	SF-12	SF-8	SF-36 V1	SF-36 V2	SF-36 V1	SF-36 V2	SF-12	SF-8
PCS	–	12	8	–	–	50 (10)	50 (10)	50 (10)	50 (10)
MCS	–	12	8	–	–	50 (10)	50 (10)	50 (10)	50 (10)
Función física	10	2	1	21	21	0-100	50 (10)	50 (10)	50 (10)
Rol físico	4	2	1	5	17	0-100	50 (10)	50 (10)	50 (10)
Dolor corporal	2	1	1	11	11	0-100	50 (10)	50 (10)	50 (10)
Salud general	5	1	1	21	21	0-100	50 (10)	50 (10)	50 (10)
Vitalidad	4	1	1	21	17	0-100	50 (10)	50 (10)	50 (10)
Función social	2	1	1	9	9	0-100	50 (10)	50 (10)	50 (10)
Rol emocional	3	2	1	4	13	0-100	50 (10)	50 (10)	50 (10)
Salud mental	5	2	1	26	21	0-100	50 (10)	50 (10)	50 (10)
Ítem de Transición de salud	1	0	0	5	5	–	–	–	–

*Las puntuaciones 0-100 tienen este rango y se interpretan como el peor y el mejor estado de salud, respectivamente; en las puntuaciones 50 (10), la población de referencia tiene una media de 50 con una desviación estándar de 10, por lo que valores superiores o inferiores a 50 indican un mejor o peor estado de salud, respectivamente, que la población de referencia.

siones del SF-36 y de las versiones reducidas está resumida en la tabla 5.

Versiones reducidas: SF-12 y SF-8

Aunque el cuestionario SF-36 puede ser cumplimentado en un período que oscila entre 5 y 10 min, en determinados contextos puede representar demasiado tiempo. La obtención de las 2 medidas sumario del SF-36 motivó el desarrollo de un cuestionario más corto que fuera capaz de reproducirlas con un número menor de ítems, el SF-12⁴³. Dado que éste podía ser autocontestado en una media de ≤ 2 min, se pretendía que su uso fuera dirigido a evaluar conceptos generales de salud física y mental en estudios en los que el SF-36 fuera demasiado largo.

El SF-12 está formado por un subconjunto de 12 ítems del SF-36 obtenidos a partir de regresión múltiple, incluidos 1-2 ítems de cada una de las 8 escalas del SF-36. La información de estos 12 ítems se utiliza para construir las medidas sumario física y mental del SF-12 (PCS-12 y MCS-12, respectivamente). Los ítems del cuestionario SF-12 explicaron más del 90% de la varianza de los índices sumario físico y mental del SF-36⁴³ en Estados Unidos y el 91% en España.

Al igual que con las medidas sumario del SF-36, se compararon los algoritmos de cálculo específicos para cada país con los algoritmos estándares, y se comprobó que había muy poca diferencia en el hecho de usar un algoritmo u otro⁴⁴. Mientras que la utilización del algoritmo estándar (basado en la población americana) facilita la comparación entre diferentes países, el algoritmo específico para el país permite la interpretación directa respecto de la población general del país en cuestión. El uso de uno u otro algoritmo debería quedar indicado explícitamente en las publicaciones resultantes.

Aunque el SF-12 conlleva una pérdida de la precisión en las puntuaciones con respecto al SF-36, para estudios de comparación de grupos numerosos, estas diferencias no son tan importantes dado que los intervalos de confianza de las medias de los grupos están muy determinados por el tamaño de la muestra. Así pues, el SF-12 ha demostrado ser una alternativa útil al SF-36 cuando se pretende medir la salud física y mental en general y el tamaño de muestra es elevado (500 individuos o más). Para estudios más pequeños o cuando se pretende estudiar una o más escalas del SF-36 por separado, es preferible utilizar el SF-36⁴⁴.

Recientemente ha sido desarrollado el cuestionario SF-8⁴⁵, una versión de 8 ítems del SF-36 que supone la última fase en la evolución de la familia de cuestionarios del SF-36. Contiene un único ítem para cada una de las escalas y permite el cálculo de un perfil de 8 escalas comparable al SF-36, y también de las 2 me-

das sumario, la física y la mental. El SF-8 fue construido para sustituir al SF-36 y el SF-12 en estudios de salud poblacionales en Estados Unidos y en el ámbito internacional. El desarrollo, la validación y las normas del nuevo SF-8 están documentados en el manual de referencia del cuestionario, del cual puede encontrarse más información en la página web del cuestionario.

El SF-6D: medida de la salud basada en preferencias

El SF-6D⁴⁶, un instrumento de medida de la salud basado en preferencias, ha sido obtenido recientemente a partir del SF-36. El instrumento permitirá llevar a cabo evaluaciones económicas en atención sanitaria a partir de bases de datos del SF-36 ya disponibles o futuras.

Limitaciones del SF-36

Es necesario señalar que el SF-36 no incluye algunos conceptos de salud importantes, como los trastornos del sueño, la función cognitiva, la función familiar o la función sexual. Ello es especialmente relevante cuando se diseñan estudios en que los pacientes sufren enfermedades o efectos secundarios de tratamientos que pueden afectar a alguna de estas escalas o conceptos de salud.

Se han descrito problemas para su aplicación en personas mayores, especialmente en ancianos hospitalizados (elevado porcentaje de pacientes que no contestan el cuestionario o dejan preguntas en blanco)⁴⁷⁻⁵⁰. En estudios con la versión española se han descrito problemas de cumplimentación, sobre todo en los grupos de edad avanzada, cuando se utilizó el formato de «matriz» para las opciones de respuesta⁵¹. Por ello, para facilitar al máximo su cumplimentación, en estudios con muestras de individuos ancianos es aconsejable utilizar letra de mayor tamaño con las preguntas impresas, cada una seguida de sus opciones de respuesta correspondientes (es decir, sin formato «matriz»). Además, siempre que fuera posible se debería revisar la cumplimentación de los cuestionarios autoadministrados en presencia del paciente para evitar el problema de la no respuesta.

Finalmente, hay que señalar que todavía no disponemos de valores de referencia para las 2 escalas de rol de la versión 2 del SF-36. Sin embargo, teniendo en cuenta las ventajas que supone respecto a la versión 1 y la similitud entre las 2 versiones americanas para dichas puntuaciones medias en la población general, resulta razonable asumir que las obtenidas en la población española con la versión 1 son el mejor estimador de que disponemos, y utilizarlas también como

valores de referencia para la versión 2 española. Tampoco hay estudios españoles que hayan estimado la mínima diferencia importante (MID)⁵², que se utiliza para el cálculo del tamaño de muestra en estudios de evaluación de tratamientos.

Conclusiones

La versión española del SF-36 es uno de los instrumentos genéricos más utilizados en el territorio nacional, tanto en estudios descriptivos que miden el impacto sobre la CVRS en distintas poblaciones de pacientes^{33,41,53-83} como para la evaluación de intervenciones terapéuticas^{35-38,83-94}. También se ha utilizado como referencia en la validación de nuevos instrumentos de medición^{56,95-103}. Estas experiencias y toda la información presentada sugieren que se puede recomendar el uso de este instrumento tanto en investigación como en la práctica clínica.

Links de interés sobre la familia de cuestionarios SF

Página web de referencia del cuestionario original: <http://www.sf-36.org/>

Página web para obtener los cuestionarios, el manual y algoritmo de cálculo específico de la versión española del cuestionario: <http://www.imim.es/imim/cas/c-QQV.htm>

Página web de la Red de investigación cooperativa para la Investigación en Resultados de Salud y Servicios Sanitarios (RED IRYSS): <http://www.rediryss.net>

Agradecimientos

Este trabajo ha sido realizado gracias a una beca del Instituto Carlos III (G03/202), Red de investigación cooperativa para la Investigación en Resultados de Salud y Servicios Sanitarios (RED IRYSS). Se recibió una ayuda adicional del Departament d'Universitats Recerca i Societat de la Informació (SGR 2001/SGR/00405) de la Generalitat de Catalunya.

Bibliografía

1. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36) (I). Conceptual framework and item selection. *Med Care*. 1992;30:473-83.
2. Ware JE. SF-36 health survey update. *Spine*. 2000;25:3130-9.
3. Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. *BMJ*. 2002;324:1417.

4. Alonso J, Prieto L, Anto JM. La versión española del SF-36 Health Survey (Cuestionario de Salud SF-36): un instrumento para la medida de los resultados clínicos. *Med Clin (Barc)*. 1995;104:771-6.
5. Ware JE Jr, Snow KK, Kosinski M, Gandek B. SF-36 Health Survey: manual and interpretation guide. Boston: New England Medical Center; 1993.
6. Aaronson NK, Acquadro C, Alonso J, Apolone G, Bucquet D, Bullinger M, et al. International Quality of Life Assessment (IQOLA) Project. *Qual Life Res*. 1992;1:349-51.
7. Bullinger M, Alonso J, Apolone G, Leplege A, Sullivan M, Wood-Dauphinee S, et al. Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. *International Quality of Life Assessment*. *J Clin Epidemiol*. 1998;51:913-23.
8. Wagner AK, Gandek B, Aaronson NK, Acquadro C, Alonso J, Apolone G, et al. Cross-cultural comparisons of the content of SF-36 translations across 10 countries: results from the IQOLA Project. *International Quality of Life Assessment*. *J Clin Epidemiol*. 1998;51:925-32.
9. Gonzalez N, Padierna A, Quintana J, Arostegui I, Horcajo M. Calidad de vida de los pacientes afectados de trastornos de la alimentación. *Gac Sanit*. 2001;15:18-24.
10. McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36) (II). Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care*. 1993;31:247-63.
11. Ware JE Jr, Gandek B, y el IQOLA Project Group. The SF-36 Health Survey: development and use in mental health research and the IQOLA project. *Int J Ment Health*. 1994;23:49-73.
12. Keller SD, Ware JE Jr, Bentler PM, Aaronson NK, Alonso J, Apolone G, et al. Use of structural equation modeling to test the construct validity of the SF-36 Health Survey in ten countries: results from the IQOLA Project. *International Quality of Life Assessment*. *J Clin Epidemiol*. 1998;51:1179-88.
13. Ware JE Jr, Kosinski M, Gandek B, Aaronson NK, Apolone G, Bech P, et al. The factor structure of the SF-36 Health Survey in 10 countries: results from the IQOLA Project. *International Quality of Life Assessment*. *J Clin Epidemiol*. 1998;51:1159-65.
14. Ware JE, Kosinski M, Keller SD. SF-36 physical and mental health summary scales: a user's manual. Boston: The Health Institute; 1994.
15. Ware JE Jr, Gandek B, Kosinski M, Aaronson NK, Apolone G, Brazier J, et al. The equivalence of SF-36 summary health scores estimated using standard and country-specific algorithms in 10 countries: results from the IQOLA Project. *International Quality of Life Assessment*. *J Clin Epidemiol*. 1998;51:1167-70.
16. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res*. 2002;11:193-205.
17. Nunnally JC, Bernstein IH. *Psychometric theory*. 3rd ed. New York: McGraw-Hill; 1994.
18. Fleiss JL. *The design and analysis of clinical experiments*. New York: John Wiley & Sons, Inc.; 1986.
19. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis*. 1987;40:171-8.
20. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care*. 1989;27 Suppl 3:178-89.
21. Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum Associates; 1988.

22. Deeks J, Higgins J, Altman D, Cochrane Statistical Methods Group. Review manager guide. Analysing and presenting results. Cochrane Library; 2000.
23. Sterne JA, Bradburn MJ, Egger M. Meta-analysis in Stata. En: Egger M, Smith GD, Altman DG, editors. Systematic reviews in health care. Meta-analysis in context. London: BMJ Publishing Group; 2004. p. 347-69.
24. Giraldo P, Pocovi M, Pérez-Calvo J, Rubio-Felix D, Giralto M. Report of the Spanish Gaucher's disease registry: clinical and genetic characteristics. *Haematologica*. 2000;85:792-9.
25. Alonso J, Prieto L, Ferrer M, Vilagut G, Broquetas JM, Roca J, et al. Testing the measurement properties of the Spanish version of the SF-36 Health Survey among male patients with chronic obstructive pulmonary disease. *Quality of Life in COPD Study Group. J Clin Epidemiol*. 1998;51:1087-94.
26. Alonso J, Regidor E, Barrio G, Prieto L, Rodríguez C, De La Fuente De Hoz L. Valores poblacionales de referencia de la versión española del Cuestionario de Salud SF-36. *Med Clin (Barc)*. 1998;111:410-6.
27. Ayuso-Mateos JL, Lasa L, Vázquez-Barquero JL, Oviedo A, Díez-Manrique JF. Measuring health status in psychiatric community surveys: internal and external validity of the Spanish version of the SF-36. *Acta Psychiatr Scand*. 1999;99:26-32.
28. Espinosa De Los Monteros MJ, Alonso J, Ancochea J, González A. Calidad de vida en asma: fiabilidad y validez del cuestionario genérico SF-36 aplicado a la población asmática de un área sanitaria. *Arch Bronconeumol*. 2002;38:4-9.
29. Failde I, Ramos I. Validity and reliability of the SF-36 Health Survey Questionnaire in patients with coronary artery disease. *J Clin Epidemiol*. 2000;53:359-65.
30. García Ordóñez MA, Mansilla Francisco JJ, Nieto Aragón E, Cereto MR, Salas Samper F, Vallejo Díaz M, et al. Calidad de vida relacionada con la salud de pacientes infectados por VIH medida por el Cuestionario de Salud SF-36. *An Med Interna*. 2001;18:74-9.
31. Gómez-Besteiro MI, Santiago-Pérez MI, Alonso-Hernández A, Valdes-Canedo F, Rebollo-Álvarez P. Validity and reliability of the SF-36 Questionnaire in patients on the waiting list for a kidney transplant and transplant patients. *Am J Nephrol*. 2004;24:346-51.
32. López-García E, Banegas JR, Graciani Pérez-Regadera A, Gutiérrez-Fisac JL, Alonso J, Rodríguez-Artalejo F. Valores de referencia de la versión española del Cuestionario de Salud SF-36 en población adulta de más de 60 años. *Med Clin (Barc)*. 2003;120:568-73.
33. Domingo-Salvany A, Lamarca R, Ferrer M, García-Aymench J, Alonso J, Felez M, et al. Health-related quality of life and mortality in male patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2002;166:680-5.
34. Lopez Revuelta K, García López FJ, De Álvaro Moreno F, Alonso J on behalf the CALVIDIA group. Perceived mental health at the start of dialysis as a predictor of morbidity and mortality in patients with end-stage renal disease (CALVIDIA Study). *Nephrol Dial Transplant*. 2004;19:2347-53.
35. Nauffal D, Doménech R, Martínez García MA, Compte L, Macian V, Perpina M. Noninvasive positive pressure home ventilation in restrictive disorders: outcome and impact on health-related quality of life. *Respir Med*. 2002;96:777-83.
36. Permanyer Miralda C, Brotons Cuixart C, Ribera Sole A, Moral Peláez I, Cascant Castello P, Alonso J, et al. Resultados clínicos y de calidad de vida de los pacientes tratados con angioplastia coronaria con balón o *stent*. Estudio multicéntrico prospectivo. *Rev Esp Cardiol*. 2001;54:597-606.
37. Permanyer Miralda C, Brotons Cuixart C, Ribera Sole A, Alonso Caballero J, Cascant Castello P, Moral Peláez I. Resultados después de cirugía coronaria: determinantes de calidad de vida relacionada con la salud postoperatoria. *Rev Esp Cardiol*. 2001;54:607-16.
38. Salinas Sánchez AS, Hernández MI, Segura MM, Lorenzo Romero JG, Virseda Rodríguez JA. The impact of benign prostatic hyperplasia surgery on patients' quality of life. *Urol Int*. 2002;68:32-7.
39. Ware JE Jr, Gandek B. Methods for testing data quality, scaling assumptions, and reliability: the IQOLA Project approach. *International Quality of Life Assessment. J Clin Epidemiol*. 1998;51:945-52.
40. Ware JE Jr, Kosinski M, Dewey JE. How to score version 2 of the SF-36 Health Survey. Lincoln RI: Quality Metric Incorporated; 2000.
41. Ruiz de Velasco I, Quintana J, Padierna J, Aróstegui I, Bernal A, Pérez Izquierdo J, et al. Validez del cuestionario de calidad de vida SF-36 como indicador de resultados de procedimientos médicos y quirúrgicos. *Rev Calidad Asistencial*. 2002;17:206-12.
42. Alonso J, Ferrer M, Gandek B, Ware JE, Aaronson NK, Mosconi P, et al. Health-related quality of life associated with chronic conditions in eight countries. Results from the International Quality of Life Assessment (IQOLA) project. *Qual Life Res*. 2004;13:283-98.
43. Ware JE Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34:220-33.
44. Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. *International Quality of Life Assessment. J Clin Epidemiol*. 1998;51:1171-8.
45. Ware JE Jr, Kosinski M, Dewey JE, Gandek B. How to score and interpret single-item health status measures: a manual for users of the SF-8™ Health Survey. Lincoln RI: Quality Metric Incorporated; 2001.
46. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ*. 2002;21:271-92.
47. Brazier JE, Walters SJ, Nicholl JP, Kohler B. Using the SF-36 and Euroqol on an elderly population. *Qual Life Res*. 1996;5:195-204.
48. Hayes V, Morris J, Wolfe C, Morgan M. The SF-36 Health Survey Questionnaire: is it suitable for use with older adults. *Age Ageing*. 1995;24:120-5.
49. O'Mahony PG, Rodgers H, Thomson RG, Dobson R, James OF. Is the SF-36 suitable for assessing health status of older stroke patients? *Age Ageing*. 1998;27:19-22.
50. Parker SG, Peet SM, Jagger C, Farhan M, Castleden CM. Measuring health status in older patients. The SF-36 in practice. *Age Ageing*. 1998;27:13-8.
51. Ferrer M, Alonso J. The use of the Short Form (SF)-36 questionnaire for older adults [letter]. *Age Ageing*. 1998;27:755-6.
52. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc*. 2002;77:371-83.
53. Argimon JM, Limon E, Vila J, Cabezas C. Health-related quality of life in carers of patients with dementia. *Fam Pract*. 2004;21:454-7.
54. Pique JM, Kulich KR, Vegazoc O, Jiménez J, Zapardiel J, Carlsson J, et al. Burden of gastroesophageal reflux disease. Evidence from a recent methodological study in Spain. *Gastroenterol Hepatol*. 2004;27:300-6.

55. Ariza-Ariza R, Hernández-Cruz B, Navarro-Sarabia F. Physical function and health-related quality of life of Spanish patients with ankylosing spondylitis. *Arthritis Rheum*. 2003;49:483-7.
56. Badia X, Díez-Pérez A, Álvarez-Sanz C, Díaz-López B, Díaz-Curiel M, Guillen F, et al. Measuring quality of life in women with vertebral fractures due to osteoporosis: a comparison of the OQLQ and QUALEFFO. *Qual Life Res*. 2001;10:307-17.
57. Baron-Esquivias G, Cayuela A, Gómez S, Aguilera A, Campos A, Fernández M, et al. Calidad de vida en los pacientes con síncope vasovagal. Influencia de parámetros clínicos. *Med Clin (Barc)*. 2003;121:245-9.
58. Bobes J, González MP, Bascaran MT, Arango C, Sáiz PA, Bousono M. Quality of life and disability in patients with obsessive-compulsive disorder. *Eur Psychiatry*. 2001;16:239-45.
59. Extremera N, Fernández-Berrocal P. Relation of perceived emotional intelligence and health-related quality of life of middle-aged women. *Psychol Rep*. 2002;91:47-59.
60. Juan J, Estiarte R, Colome E, Artes M, Jiménez FJ, Alonso J. Burden of illness of Crohn's disease in Spain. *Dig Liver Dis*. 2003;35:853-61.
61. López-García E, Banegas BE, Gutiérrez-Fisac JL, Pérez-Regadera AG, Ganan LD, Rodríguez-Artalejo F. Relation between body weight and health-related quality of life among the elderly in Spain. *Int J Obes Relat Metab Disord*. 2003;27:701-9.
62. Mena-Martin FJ, Martin-Escudero JC, Simal-Blanco F, Carretero-Ares JL, Arzuza-Mouronte D, Herreros-Fernández V. Health-related quality of life of subjects with known and unknown hypertension: results from the population-based Hortega study. *J Hypertens*. 2003;21:1283-9.
63. Naves Díaz M, Díaz López JB, Rodríguez Rebollos A, Gómez Alonso C, Díaz Corte C, Cannata Andia J. Efecto de la fractura vertebral sobre la calidad de vida relacionada con la salud en población asturiana mayor de 54 años. *Med Clin (Barc)*. 2001;116:533-5.
64. Padierna A, Quintana JM, Arostegui I, González N, Horcajo MJ. The health-related quality of life in eating disorders. *Qual Life Res*. 2000;9:667-74.
65. Rebollo P, Bobes J, González MP, Sáiz P, Ortega F. Factores asociados a la calidad de vida relacionada con la salud (CVRS) de los pacientes en terapia renal sustitutiva (TRS). *Nefrología*. 2000;20:171-81.
66. Rebollo P, Ortega F, Baltar JM, Álvarez-Ude F, Álvarez Navascues R, Álvarez -Grande J. Is the loss of health-related quality of life during renal replacement therapy lower in elderly patients than in younger patients? *Nephrol Dial Transplant*. 2001;16:1675-80.
67. Salinas-Sánchez AS, Hernández-Millan I, Lorenzo-Romero JG, Segura-Martin M, Fernández-Olano C, Virseda-Rodríguez JA. Quality of life of patients on the waiting list for benign prostatic hyperplasia surgery. *Qual Life Res*. 2001;10:543-53.
68. Sánchez-Cruz JJ, Cabrera-Leon A, Martín-Morales A, Fernández A, Burgos R, Rejas J. Male erectile dysfunction and health-related quality of life. *Eur Urol*. 2003;44:245-53.
69. Tusell JM, Aznar JA, Querol F, Quintana M, Moreno M, Gorina E. Results of an orthopaedic survey in young patients with severe haemophilia in Spain. *Haemophilia*. 2002;8 Suppl 2:38-42.
70. Gene J, Moliner C, Contel JC, Tintore I, Villafañila R, Schorlemmer C. Salud y utilización de servicios en anciano, según el nivel de convivencia. *Gac Sanit*. 1997;11:214-20.
71. Regidor E, Barrio G, De la Fuente L, Domingo A, Rodríguez C, Alonso J. Association between educational level and health related quality of life in Spanish adults. *J Epidemiol Community Health*. 1999;53:75-82.
72. Badia X, Mearin F, Balboa A, Baro E, Caldwell E, Cucala M, et al. Burden of illness in irritable bowel syndrome comparing Rome I and Rome II criteria. *Pharmacoeconomics*. 2002;20:749-58.
73. Buck D, Jacoby A, Baker GA, Ley H, Steen N. Cross-cultural differences in health-related quality of life of people with epilepsy: findings from a European study. *Qual Life Res*. 1999;8:675-85.
74. Rebollo P, Ortega F, Baltar JM, Badia X, Álvarez-Ude F, Díaz-Corte C, et al. Health related quality of life (HRQOL) of kidney transplanted patients: variables that influence it. *Clin Transplant*. 2000;14:199-207.
75. Rebollo P, González MP, Bobes J, Sáiz P, Ortega F. Interpretación de los resultados de la calidad de vida relacionada con la salud de pacientes en terapia sustitutiva de la insuficiencia renal terminal. *Nefrología*. 2000;20:431-9.
76. Dowrick C, Dunn G, Ayuso-Mateos JL, Dalgard OS, Page H, Lehtinen V, et al. Problem solving treatment and group psychoeducation for depression: multicentre randomised controlled trial. *Outcomes of Depression International Network (ODIN) Group*. *BMJ*. 2000;321:1450-4.
77. Gamiz MJ, López-Escámez JA. Health-related quality of life in patients over sixty years old with benign paroxysmal positional vertigo. *Gerontology*. 2004;50:82-6.
78. Guitera V, Muñoz P, Castillo J, Pascual J. Quality of life in chronic daily headache: a study in a general population. *Neurology*. 2002;58:1062-5.
79. Mones J, Adan A, Segú JL, López JS, Artes M, Guerrero T. Quality of life in functional dyspepsia. *Dig Dis Sci*. 2002;47:20-6.
80. Monzon MJ, Láinez MJ. Quality of life in migraine and chronic daily headache patients. *Cephalalgia*. 1998;18:638-43.
81. Quintana JM, Arostegui I, Cabriada J, López I, Perdigo L. Predictors of improvement in health-related quality of life in patients undergoing cholecystectomy. *Br J Surg*. 2003;90:1549-55.
82. Rebollo P, Ortega F, Valdes C, Fernández-Vega F, Ortega T, García-Mendoza M, et al. Influence of erectile dysfunction on health related quality of life of male kidney transplant patients. *Int J Impot Res*. 2004;16:282-7.
83. Valderrabano F. Quality of life benefits of early anaemia treatment. *Nephrol Dial Transplant*. 2000;15 Suppl 3:23-8.
84. Targarona EM, Novell J, Vela S, Cerdan G, Bendahan G, Torrubia S, et al. Mid term analysis of safety and quality of life after the laparoscopic repair of paraesophageal hiatal hernia. *Surg Endosc*. 2004;18:1045-50.
85. Colas R, Muñoz P, Temprano R, Gómez C, Pascual J. Chronic daily headache with analgesic overuse: epidemiology and impact on quality of life. *Neurology*. 2004;62:1338-42.
86. Gual A, Balcells M, Torres M, Madrigal M, Díez T, Serrano L. Sertraline for the prevention of relapse in detoxicated alcohol dependent patients with a comorbid depressive disorder: a randomized controlled trial. *Alcohol Alcohol*. 2003;38:619-25.
87. Rullan M, Cerda L, Frontera G, Llobera J, Masmiquel L, Olea JL. Ensayo clínico, controlado con placebo, triple ciego, para evaluar la eficacia de una heparina de bajo peso molecular (bemiparina) en el tratamiento de las úlceras tórpidas del pie diabético, en atención primaria. *Aten Primaria*. 2003;31:539-44.
88. Brotons Cuixart C, Moral Peláez I, Permanyer Miralda G, Ribera Sole A, Cascant Castello P. Control terapéutico de los factores de riesgo y calidad de vida en los pacientes trata-

- dos mediante derivación aortocoronaria. *Med Clin (Barc)*. 2001;116:241-5.
89. Montserrat JM, Ferrer M, Hernández L, Farre R, Vilagut G, Navajas D, et al. Effectiveness of CPAP treatment in daytime function in sleep apnea syndrome: a randomized controlled study with an optimized placebo. *Am J Respir Crit Care Med*. 2001;164:608-13.
 90. Padierna A, Quintana JM, Arostegui I, González N, Horcajo MJ. Changes in health related quality of life among patients treated for eating disorders. *Qual Life Res*. 2002;11:545-52.
 91. Permanyer-Miralda G, Brotons Cuixart C, Ribera Sole A, Cascant Castello P, Moral Peláez I, Pons JM, et al. Desigual perfil clínico, calidad de vida y mortalidad hospitalaria en pacientes operados de injerto aortocoronario en centros públicos y privados de Catalunya. The CIRCORCA Study. *Rev Esp Cardiol*. 1998;51:806-15.
 92. Aiarzaguena JM, Grandes G, Alonso-Arbiol I, Campo Chavala JL, Oleaga Fernández MB, Marco DJ. Abordaje biopsicosocial de los pacientes somatizadores en las consultas de atención primaria un estudio piloto. *Aten Primaria*. 2002;29:558-61.
 93. Hernández-Osma E, Cairols MA, Martí X, Barjau E, Riera S. Impact of treatment on the quality of life in patients with critical limb ischaemia. *Eur J Vasc Endovasc Surg*. 2002;23:491-4.
 94. Quintana JM, Cabriada J, Arostegui I, De Tejada IL, Bilbao A. Quality-of-life outcomes with laparoscopic vs open cholecystectomy. *Surg Endosc*. 2003;17:1129-34.
 95. Badia Llach X, Castro Díaz D, Conejero Sugranes J. Validez del cuestionario King's Health para la evaluación de la calidad de vida en pacientes con incontinencia urinaria. *Med Clin (Barc)*. 2000;114:647-52.
 96. Escobar A, Quintana JM, Bilbao A, Azkarate J, Guenaga JI. Validation of the Spanish version of the WOMAC questionnaire for patients with hip or knee osteoarthritis. *Western Ontario and McMaster Universities Osteoarthritis Index*. *Clin Rheumatol*. 2002;21:466-71.
 97. Escriba-Aguir V, Mas Pons R, Flores Reus E. Validación del Job Content Questionnaire en personal de enfermería hospitalario. *Gac Sanit*. 2001;15:142-9.
 98. Martínez Martín P, Frades B, Jiménez Jiménez FJ, Pondal M, López Lozano JJ, Vela L, et al. La versión española del PDQ-39: fiabilidad y correlación con Short Form Health Survey (SF-36). *Neurología*. 1999;14:159-63.
 99. Quintana JM, Cabriada J, López de Tejada I, Varona M, Oribe V, Barrios B, et al. Traducción y validación del Índice de Calidad de Vida Gastrointestinal (GIQLI). *Rev Esp Enferm Dig*. 2001;93:693-706.
 100. Rebollo P, Ortega F, Ortega T, Valdes C, García-Mendoza M, Gómez E. Spanish validation of the Kidney Transplant Questionnaire: a useful instrument for assessing health related quality of life in kidney transplant patients. *Health Qual Life Outcomes*. 2003;1:56.
 101. Torrens C, Orient F, Vila F, Escalada F, Marín M. Análisis comparativo entre la escala de constant y el cuestionario de salud SF-36 en pacientes con patología subacromial. *Rev Ortop Traumatol*. 2000;44:447-51.
 102. Badia X, Prieto L, Roset M, Díez-Pérez A, Herdman M. Development of a short osteoporosis quality of life questionnaire by equating items from two existing instruments. *J Clin Epidemiol*. 2002;55:32-40.
 103. Escriba V, Mas R, Cárdenas M, Pérez S. Validación de la escala de estresores laborales en personal de enfermería: «the nursing stress scale». *Gac Sanit*. 1999;13:191-200.