

**ESTIMACIÓN DE UN ALGORITMO PARA EL CÁLCULO DE
UTILIDADES DE ESTADOS DE SALUD DESCRITOS MEDIANTE
EL SF-6D(SF-36) (*)**

Junio 2008

Equipo investigador

Fernando Ignacio Sánchez Martínez

José María Abellán Perpiñán

Jorge Eduardo Martínez Pérez

Ildfonso Méndez Martínez

Grupo de Investigación en Economía de la salud y evaluación económica

Departamento de Economía Aplicada

Universidad de Murcia



(*) Informe correspondiente al proyecto *Indicadores de calidad de vida relacionada con la salud y su aplicación a la evaluación de tecnologías sanitarias*, desarrollado en el marco del Convenio de Colaboración suscrito con fecha 22 de diciembre de 2006 entre la C.A.R.M.-Consejería de Sanidad y la Universidad de Murcia, en materia de Economía de la Salud, - BORM N° 23, 29/01/2007.

ÍNDICE

1. Introducción.....	3
2. Revisión bibliográfica	7
2.1. Objetivos y métodos de búsqueda	7
2.2. Resultados de la revisión	9
2.2.1 Medidas de calidad de vida relacionadas con la salud	9
2.2.2. Perfiles de salud.....	11
2.2.3. Propiedades empíricas del SF-36 en comparación a otras medidas	17
2.2.4. Medidas basadas en preferencias.....	20
2.2.5. Rasgos básicos del SF-6D	25
2.2.6. Modelizaciones de las utilidades SF-6D	32
2.2.7. Propiedades empíricas del SF-6D en comparación a otras medidas	47
3. Motivación del estudio.	53
3.1. Antecedentes.....	53
3.2. Objetivos.....	54
4. Material y métodos.	57
4.1. Selección de estados y métodos de valoración.....	57
4.2. Recogida de datos	60
4.2.1. La muestra	60
4.2.2. El cuestionario	61
4.2.3. Cálculo de las utilidades	63
4.2.4. Cálculo de las puntuaciones resumen del SF-36	63
4.3. Métodos de estimación de la tarifa SF-6D	66
4.3.1. Modelos de efectos aleatorios.....	66
4.3. 2. Enfoque semiparamétrico	67
5. Resultados.....	73
5.1. Participantes en el estudio	73
5.2. Utilidades de los estados de salud	75
5.3. Modelización de la tarifa SF-6D	78
5.3.1. Modelos de efectos aleatorios.....	78
5.3.2. Estimaciones semiparamétricas	83
6. Conclusiones.....	87
6.1. Recapitulación de resultados	87
6.1.1. SF-36	87
6.1.2. Utilidades de los estados de salud SF-6D.....	88
6.1.3. Modelos para predecir la tarifa SF-6D	89
6.2. Aplicaciones del algoritmo estimado	89
Referencias	91
Apéndice 1. El cuestionario.....	101
Introducción: descripción del sistema SF-6D.....	101
Puntuación de estados en la escala visual.....	101
Obtención de medidas directas de utilidad (PLE)	102
Autoevaluación del propio estado de salud	105
Características socio-demográficas	106
Procedimiento para la obtención de los valores de indiferencia (PEST)	107
Apéndice 2. Resultados del SF-36.....	109
Descripción de la muestra.....	109
Análisis de las puntuaciones SF-36.....	111
Apéndice 3. Ejemplo de aplicación del algoritmo SF-6D.....	141

1. Introducción

El interés por la realización del estudio que describimos en las páginas siguientes fue manifiesto por las dos partes firmantes del Convenio, conviniendo todos los actores involucrados en su financiación, diseño y ejecución ulterior la gran relevancia que posee para potenciales aplicaciones prácticas presentes y futuras. Sin afán por ser exhaustivos, un primer resultado tangible extraído de este proyecto es la obtención de una “radiografía” (tentativa si se quiere) de la calidad de vida relacionada con la salud (CVRS) en la Región de Murcia. El instrumento empleado para tal fin ha sido el cuestionario Short Form-36, mejor conocido por sus siglas SF-36. Que sepamos, es la primera ocasión que se utiliza dicha herramienta en la Región de Murcia con el objeto de aproximarse a la CVRS de la población general, no de grupos específicos de pacientes. A esto hay que añadir que el cuestionario SF-36 ha sido administrado mediante entrevistas personales, no postales o telefónicas, con la ayuda de una interfaz informática concebida por el equipo investigador. Las garantías de fiabilidad y validez son pues elevadas.

Una segunda consecuencia del trabajo realizado es la obtención directa, para la muestra entrevistada, de indicadores de bienestar o utilidad (también conocidos como “pesos” de calidad de vida) asociados a 78 estados de salud SF-6D. Como se explica a lo largo de la Sección 2 de esta memoria, el sistema de utilidad multiatributo SF-6D es una medida de CVRS basada en preferencias, diferente por tanto de instrumentos como el ya citado SF-36. Ocurre, sin embargo, que uno y otro guardan una estrecha relación, ya que el sistema SF-6D caracteriza un problema o estado de salud mediante una combinación de atributos (o dimensiones) y niveles de gravedad obtenidos como el resultado de “redimensionar” las casi infinitas combinaciones que propicia el SF-36. Por utilizar un símil muy sencillo, podríamos decir que tanto el SF-36 como el SF-6D son mapas de la realidad (la salud), pero que difieren en la escala que emplean para describir esa misma realidad. De nuevo, hasta donde llega el conocimiento del equipo investigador, la medición de utilidades SF-6D es una novedad en el contexto de la sanidad murciana. Asimismo, y en esto la novedad es de alcance superior al regional, el método de obtención de preferencias que se presenta es inédito en el panorama de los estudios con el SF-6D. Los antecedentes mostrados en la Sección 3 revelan que el único método de valoración que se ha empleado hasta la fecha ha sido el de la “lotería o juego estándar”

(*standard gamble*). Técnica ésta, bien es cierto, de honda raigambre en el acervo metodológico de la Economía de la Salud (en particular de la evaluación económica de las intervenciones sanitarias), pero técnica también cada vez más cuestionada a medida que nuevas investigaciones han ido acumulando evidencia en su contra, consecuencia de sesgos como el denominado *efecto certeza* y, más generalmente, la *distorsión de las probabilidades*. Por otro lado, justificaciones pretéritas basando la supremacía de este método en su conexión con la teoría de la utilidad esperada (o teoría von Neumann-Morgenstern), han perdido fuerza conforme se han falseado empíricamente algunos de los supuestos fundamentales de dicha teoría (singularmente el llamado “axioma de independencia de alternativas irrelevantes”). Una extensa corriente de investigación psicológica y económica ha hallado evidencia suficiente como para poder afirmar sin temor a errar que la validez descriptiva de la teoría de la utilidad esperada no se sostiene (al menos, no en todos los contextos de decisión). De ahí el interés que posee ensayar nuevas fórmulas de medición de las preferencias, no expuestas a priori a este tipo de debilidades. La elección del equipo investigador recayó en el método bautizado por Bleichrodt, Abellán, Pinto y Méndez (2007) como *probability lottery equivalence*. Investigaciones previas sugieren que este método es teóricamente más robusto que la lotería estándar (la probabilidad de que la teoría de la utilidad esperada se observe es mayor), y que parece también más consistente en términos empíricos (describe mejor las preferencias). No obstante, la agenda investigadora en torno a este tipo de procedimientos acaba de abrirse en el terreno de la Economía de la Salud. Por nuestra parte podemos decir que los resultados ofrecidos en la Sección 5 sugieren que, como poco, es una técnica factible, perfectamente aplicable en estudios de índole similar al que nos ocupa aquí.

Por último, presentamos en esta memoria una estimación de varios modelos que permiten predecir, sobre la base de las utilidades obtenidas de modo directo para la muestra encuestada, un peso de calidad de vida para cada uno de los 18000 estados SF-6D posibles (este conjunto de pesos predichos suele referirse como “tarifa”). Muy pocos estudios han realizado estimaciones de este tenor (tan sólo tres han sido publicados que sepamos), y únicamente para el Reino Unido. Constituyen las estimaciones efectuadas, por tanto, una primicia en el ámbito español. Como antes, sucede además que algunas de esas estimaciones son también inéditas en el panorama internacional. El apartado 5.3.2. muestra el resultado de aplicar una suerte de ajuste semiparamétrico a los modelos de efectos aleatorios convencionales, de manera que se garantiza la

Introducción

representatividad de cada una de las submuestras en que se dividió la muestra total. Este diseño entre-muestras fue adoptado con vistas a evitar lo máximo posible sesgos y errores causados por el cansancio, el aburrimiento o, en definitiva, la sobrecarga cognitiva que representaría valorar individualmente un número excesivo de estados. No hay que olvidar que son seis las dimensiones recreadas mediante el sistema SF-6D, no todas ellas con el mismo número de niveles; muchos atributos a los que prestar atención, lo cual aconseja reducir lo más posible el número de tareas acometidas por cada entrevistado. Por lo demás, el diseño entre-muestras es habitual en la literatura especializada, no sólo en el caso del SF-6D, sino también de otros instrumentos afines; el EQ-5D, por ejemplo. Se deriva de esto que la innovación metodológica abordada en este proyecto es una fuente potencial de mejora para muchos otros estudios: en general, todos aquellos que hayan estimado tarifas sociales recurriendo a una selección de estados de salud no valorados en su integridad por cada sujeto de la muestra.

Comenzábamos esta introducción recordando la confluencia de pareceres entre la D.G. de Calidad y el equipo investigador al respecto de la relevancia social que podía tener este proyecto. Hay, como entonces mencionábamos, varias posibles aplicaciones prácticas de los resultados que, a grandes rasgos, hemos sintetizado en los párrafos precedentes. Las Conclusiones de esta memoria ofrecen no sólo una recapitulación de sus contenidos, sino también sugieren aquellas propuestas que, a juicio de los integrantes del equipo investigador, mejor pueden explotar las potencialidades que subyacen a los logros reunidos en estas páginas. La última palabra, claro está, la tendrá el decisor político, gestor legítimo de los recursos públicos. Confiamos en que nuestras sugerencias inviten cuando menos a la reflexión que precede a la toma de decisiones.

2. Revisión bibliográfica

2.1. *Objetivos y métodos de búsqueda*

El objetivo fundamental de esta revisión es contribuir al mejor conocimiento de uno de los instrumentos de medición de calidad de vida relacionada con la salud (CVRS) de última generación: el denominado SF-6D. El interés que suscita esta herramienta está suficientemente justificado, toda vez que permite atribuir utilidades (índices de preferencias) a una vasta colección de estados de salud contruidos a partir de las conocidas medidas multiatributo SF-36 y SF-12.

La revisión efectuada fue de tipo metodológico, esto es, no se pretendió cubrir todas las publicaciones relacionadas con el SF-6D (p.ej. aplicaciones a enfermedades específicas), sino más bien aquellas concernientes a los fundamentos metodológicos de la medida, la estimación del modelo que permite ligar las utilidades a la descripción de los estados de salud, y la comparación de dichas utilidades con las obtenidas mediante otras medidas basadas en preferencias anteriores al SF-6D (primordialmente el EQ-5D).

Los criterios de búsqueda, revistas consultadas y recursos de internet empleados se describen en detalle a continuación:

- *Criterios de búsqueda.*

Artículos en inglés o castellano que aborden cuestiones metodológicas y/o conceptuales acerca de la relación entre el SF-36 y/o el SF-12 y/o el SF6D y la obtención de utilidades, publicados desde 1998 (fecha de aparición del primer artículo sobre el SF-6D) hasta el presente. En particular se han utilizado como criterios de inclusión los siguientes temas:

- Modelización del algoritmo que permite convertir puntuaciones en las escalas SF-36 y/o SF-12 en utilidades de estados de salud (descritos mediante el SF6D o cualquier otro sistema como, por ejemplo, el EQ-5D).
- Procedimientos de “mapeo” (mapping) que permitan pasar del SF-36 y/o el SF-12 al SF6D o cualquier otro sistema descriptivo multiatributo (p.ej. el mencionado EQ-5D).

- Ejercicios de validez, fiabilidad y sensibilidad al cambio del SF6D, así como comparaciones entre las utilidades de estados descritos mediante este sistema y cualquier otro (p.ej. EQ-5D).
- ***Palabras clave.***
 - (SF-36 OR Short Form OR SF36 OR SF-12 OR SF12 OR SF-6D OR SF6D) and (utility OR utilities OR preference based measurement OR preference assessment OR preference scores OR quality weights OR algorithm OR preference OR tariffs OR scoring formula OR multiattribute health status system OR multi-attribute health status system)
 - (SF-36 OR Short Form-36 OR SF-12 OR Short Form-12) and (EuroQol OR Euroqol OR EQ-5D)
 - (SF-36 OR Short Form-36 OR SF-12 OR Short Form-12) and (Health Utility Index OR HUI)
 - (SF-36 OR Short Form-36 OR SF-12 OR Short Form-12) and (Quality of Well-Being OR QWB)
- ***Publicaciones en las que se ha buscado.***
 - Bases de datos: *EconLit, MEDLINE, ISI web of knowledge.*
 - Revistas: *British Medical Journal, European Journal of Health Economics, European Journal of Public Health, Gaceta Sanitaria, Health Affairs , Health and Quality of Life Outcomes, Health Economics, Health Economics in Prevention and Care, Health Expectations, Health Policy, International Journal of Technology Assessment in Health Care, International Journal on Quality in Health Care, Journal of Health Economics, Journal of Public Health Medicine, Lancet , Medical Care, Medical Decision Making, Medicina Clínica, Pharmacoeconomics, Quality of Life Research, Social Science and Medicine, Value in Health.*
 - Editoriales: *Blackwell Sinergy, ScienceDirect, Wiley InterScience.*
 - Literatura gris: *Health Technology Assessment reports.*
 - Buscadores: www.google.com, <http://scholar.google.es/>.

2.2. Resultados de la revisión

2.2.1 Medidas de calidad de vida relacionadas con la salud

La calidad de vida relacionada con la salud (CVRS) es un concepto que idealmente abarcaría todos y cada uno de los atributos o dimensiones que potencialmente puedan caracterizar a la salud. Esto requiere considerar tanto el dominio de la función física (capacidad para ejercitar actividades físicas diarias, desplazamientos y cuidado personal), como el plano psicológico (ansiedad, depresión, bienestar emocional), hasta incluso el rol social.

Los primeros intentos por medir la CVRS se remontan a la década de 1940 (Karnofsky y Burchenal, 1949), si bien no es hasta los años sesenta-setenta cuando podemos encontrar lo que hoy en día se concibe como medidas de CVRS, esto es, medidas caracterizadas por tres rasgos fundamentales (Badía y Lizán, 2003):

- La importancia de los dominios físico, psíquico y social en la multidimensionalidad del concepto.
- El carácter subjetivo de la medición, efectuada a partir de las respuestas de los pacientes a las preguntas o ítems en que se descompone cada una de las dimensiones de la CVRS.
- La conversión de las respuestas a los ítems en puntuaciones numéricas que se agregan para obtener puntuaciones parciales y/o globales.

Tras el concepto de CVRS y los rasgos básicos que acabamos de enunciar, se esconden no obstante dos aproximaciones muy diferentes cuando se afronta la tarea de su medición. En general, podemos afirmar que se pueden seguir dos enfoques provenientes de tradiciones científicas distintas: el enfoque de la Psicología (el enfoque *psicométrico* podríamos decir) y el enfoque propio de la Economía de la Salud (o simplemente *económico*). Los dos enfoques tienen en común su fundamento en la *percepción subjetiva* del paciente. Ambos difieren, no obstante, en sus objetivos y en los medios que emplean. Las medidas o escalas de CVRS del enfoque psicométrico pretenden reflejar el cambio en la salud experimentado por un paciente como efecto del tratamiento, o pretenden discriminar entre pacientes con diferentes niveles de gravedad, utilizando para ello cuestionarios en los que las respuestas se miden en escalas ordinales

tipo Likert¹. En cambio, las *medidas basadas en las preferencias* del enfoque económico –también denominadas escalas de utilidad multiatributo– intentan medir la intensidad con que el paciente prefiere un estado de salud a otro, empleando para tal finalidad escalas cardinales definidas por los puntos 0 y 1 correspondientes a la muerte y la salud plena, respectivamente.

Dentro del primer grupo de medidas (las escalas de CVRS no basadas en las preferencias) resulta habitual efectuar una división según la especificidad del instrumento en cuestión. Así, podemos distinguir entre medidas *específicas* y medidas *genéricas* (o perfiles de salud). Las primeras pretenden medir la CVRS asociada a pacientes afectados por una patología específica –asma, diabetes, depresión, etc.- de manera que la propiedad psicométrica de sensibilidad al cambio² esté asegurada. Ejemplos de tales medidas específicas son, entre otras, *Ashma Quality of Life Questionnaire*, *Skindex-29* o *Diabetes Quality of Life Questionnaire*. Si bien estos cuestionarios presentan la ventaja de su mayor sensibilidad, a cambio resulta harto difícil extrapolar sus ítems a otras enfermedades diferentes. Esto es, un cuestionario ideado para capturar la CVRS de los diabéticos sólo será válido, fiable y factible³ con ese tipo de enfermos, no con otros. Como es lógico, esto plantea un problema considerable cuando se aspira a cartografiar la CVRS de la población general, y no de un grupo específico de pacientes. Por esta razón (aunque no sólo) se dispone de medidas genéricas de CVRS, no vinculadas a enfermedades concretas, sino susceptibles de describir el estado de salud de poblaciones diversas⁴. Algunas de las medidas genéricas más importantes son *Nottingham Health Profile*, *Short Form-36*, o *Sickness Impact Profile*. Cabe esperar que su sensibilidad al cambio sea inferior a la correspondiente a las medidas específicas, sin embargo, su mayor ductilidad las convierte en una

¹ Un ejemplo muy sencillo de escala tipo Likert nos lo ofrece la pregunta acerca del estado de salud autopercebido incluida en la Encuesta Nacional de Salud que elabora el INE, la cual ofrece cinco posibles respuestas de carácter categórico: muy malo, malo, regular, bueno, muy bueno.

² Por dicha propiedad se hace referencia a la capacidad de la escala para detectar cambios en la salud de los pacientes a consecuencia de la aplicación de un tratamiento.

³ En general, la propiedad de validez requiere que el instrumento mida realmente aquello para lo que fue concebido. La fiabilidad por su parte remite a la precisión con que se efectúan las mediciones. A su vez la factibilidad o viabilidad implica que la administración del instrumento de medición sea cognitivamente sencilla, de modo que se comprendan fácilmente las preguntas y no se tarde demasiado tiempo en responderlas. Una explicación más extensa de este tipo de propiedades y su medición se aborda en el epígrafe 2.2.3.

⁴ Las medidas específicas suelen utilizarse en contextos *ad hoc* (ensayos clínicos), lo cual impide la comparabilidad de los registros obtenidos. Por tanto, las medidas genéricas no son sólo atractivas para los economistas, sino también para los epidemiólogos y los clínicos en general.

herramienta muy valiosa cuando lo que interesa es contar con un criterio homogéneo (para todos los individuos igual) de clasificación del estado de salud.

Las medidas de CVRS basadas en preferencias se caracterizan por constar de dos elementos: un sistema descriptivo multiatributo de carácter genérico, y una regla de valoración o algoritmo que sirve para generar los índices cardinales de bienestar (intensidad de preferencias) o utilidades. El primero de estos dos componentes es análogo en su naturaleza a cualquiera de las medidas genéricas de CVRS que hemos citado antes. Las únicas diferencias afloran en cuanto a las dimensiones y niveles que integran, pero no en la estructura general que se utiliza para describir el estado de salud. Por tanto, su rasgo distintivo radica en el mencionado algoritmo que acompaña al componente meramente descriptivo. Algunos de los ejemplos más conocidos de este tipo de instrumentos son el *EuroQol (EQ)-5D*, el *SF-6D* y el *Health Utility Index (HUI)*.

Como veremos en el apartado 2.2.5, el SF-6D nace como un intento de intermediar entre una medida genérica de CVRS preexistente (el SF-36 primero y, posteriormente, el SF-12) y el algoritmo generador de utilidades que todo instrumento basado en preferencias posee. La principal motivación para tender ese “puente”, responde al interés que supone contar con una medida basada en preferencias que conserve parte de la riqueza informativa albergada en los cuestionarios *Short Form*. Así pues, destinaremos las páginas restantes de esta sección a mostrar un resumen de las medidas genéricas de CVRS más importantes, haciendo especial referencia al SF-36, intentado también ubicar al SF-6D dentro del elenco de las medidas basadas en preferencias.

2.2.2. Perfiles de salud

Los perfiles de salud no están concebidos para ofrecer una medida resumen del bienestar o utilidad del individuo. Lo que pretenden es caracterizar la CVRS mediante una estandarización de la misma sobre la base de una determinada combinación de dimensiones e ítems. Sadana (2002) ofrece una crónica de la evolución experimentada por este tipo de medidas desde los años sesenta hasta la actualidad. De acuerdo a este autor, los primeros perfiles de salud fueron concebidos como un medio para describir estados de salud graves, relacionados en particular con aquellas personas que requieren cuidados de larga duración. Un ejemplo de medida perteneciente a esta primera generación es la medida conocida como *Activities of Daily Living* (Katz et al., 1963).

Las medidas desarrolladas a partir de los años setenta, en cambio, fueron diseñadas para reflejar la salud de la población general, ampliándose notablemente las dimensiones cubiertas por los ítems de los cuestionarios. Cabe ubicar en esta segunda oleada de instrumentos aquellos aplicados con mayor frecuencia en los estudios sobre CVRS: *Sickness Impact Profile (SIP)* (Bergner et al., 1976, 1981) y *SIP68* (de Bruin et al., 1994; Nanda et al., 2003); *Nottingham Health Profile (NHP)* (Hunt et al., 1985, 1986); *COOP Charts* (Nelson et al., 1987; Nelson et al., 1996) y *COOP/WONCA Charts* (Van Weel, 1993); *Short-Form 36 (SF-36)* (Brazier et al., 1992; Jenkinson 1996a, 1996b; Ware and Sherbourne, 1992; Ware et al., 1993, 1994); y el *Short-Form 12 (SF-12)* (Ware et al., 1995a; Ware et al., 1995b; Ware et al., 1996). Procedemos a continuación a describir brevemente cada una de estas medidas.

Tabla 1. Perfiles de salud más importantes

Sistema descriptivo	Procedencia	Atributos	Ítems	Traducción al castellano validada *
SIP ^a	EE.UU.	12	136	Badía y Alonso (1994) ^b
NHP	Reino Unido	6	38	Alonso et al. (1990) ^c
COOP/-WONCA ^d	EE.UU.	6(+1 op.)	6(+1 op.)	Lizán y Reig (1999)
SF-36 ^e	EE.UU.	8	36	Alonso et al. (1995)
SF-12 ^f	EE.UU.	8	12	No

- a. Existe una versión abreviada conocida como SIP68 (de Bruin et al., 1994).
- b. El nombre dado en la traducción al SIP es el de *Perfil de las Consecuencias de la Enfermedad*.
- c. El nombre dado en la traducción al NHP es el de *Perfil de Salud de Nottingham*.
- d. Las viñetas COOP/WONCA tienen 6 dimensiones, más una séptima opcional. Las viñetas COOP originales poseen 9 dimensiones con 1 ítem por dimensión.
- e. Existen dos versiones del cuestionario SF-36: SF-36 v1.0 (Ware y Sherbourne, 1992) y SF-36 v2.0 (Ware et al., 2000).
- f. Existen dos versiones del cuestionario SF-12: SF-12 v1.0 (Ware et al., 1996) y SF-12 v2.0 (Ware et al., 2002).

*: Por traducción validada se entiende aquella que se ha obtenido merced a un protocolo de adaptación transcultural del instrumento de CVRS original para su utilización, con las mismas garantías psicométricas, en un ámbito cultural y lingüístico diferente. Dicha adaptación se consigue (Guillemin et al., 1993) mediante una serie de traducciones y retrotraducciones a partir de la versión original del instrumento, para evaluar finalmente el cuestionario traducido mediante un estudio piloto (*pre-test*).

Sickness Impact Profile/ SIP-68

El SIP es un cuestionario estandarizado concebido para informar acerca de los cambios acaecidos en la conducta de las personas a consecuencia de padecer una enfermedad. Consta de 136 preguntas de carácter binario (sí/no), agrupadas en 12 dimensiones o categorías (véase la Tabla 2). Tres de estas categorías (desplazamientos, movilidad y cuidado personal) pueden agregarse dando lugar a una puntuación resumen de la dimensión salud física. Asimismo puede obtenerse una segunda puntuación resumen de

la dimensión psicosocial mediante la agregación de las puntuaciones de otros cuatro atributos (comunicación, interacción social, atención y estado emocional). Las cinco dimensiones restantes se tratan de modo independiente, sin originar una nueva medida resumen. Finalmente cabe estimar un índice global, de modo que cuanto mayor sea su valor tanto peor es la salud del sujeto. Las puntuaciones se calculan dividiendo la suma de una serie de pesos predeterminados ligados a los ítems entre la máxima puntuación posible, dando lugar a un rango de puntuaciones 0-100.

El SIP68 es una versión reducida del cuestionario SIP original, construida a partir de un análisis de componentes principales. Consta de 68 preguntas (en lugar de las 136 originales), agrupadas en 6 escalas (autonomía somática, control de la movilidad, autonomía psicológica y comunicación, comportamiento social, estabilidad emocional, y amplitud de la movilidad). Esta medida permite computar puntuaciones resumen para tres grandes dimensiones (física, psicológica y social), así como calcular una puntuación global. Todas las puntuaciones se obtienen como la suma no ponderada de los diferentes ítems. La puntuación global abarca un intervalo de valores comprendido entre 0 (mejor salud posible) y 68 (peor salud posible). Las puntuaciones resumen poseen un máximo diferente cada una de ellas, dependiendo del número de ítems incluidos en la escala (p.ej. la puntuación resumen de la dimensión física abarca un rango de valores comprendidos entre 0 y 29).

Nottingham Health Profile

El NHP es un cuestionario auto-administrado que consta de dos partes, diseñado inicialmente para su uso en el ámbito de la atención primaria. La primera parte posee 38 ítems, correspondientes a seis dimensiones: energía (3 ítems), dolor (8 ítems), movilidad física (8 ítems), reacciones emocionales (9 ítems), sueño (5 ítems) y aislamiento social (5 ítems). La segunda parte del cuestionario consiste en siete preguntas acerca de la existencia de limitaciones impuestas por los problemas de salud para la realización de actividades cotidianas. Las opciones de respuesta de todos los ítems son dicotómicas (sí/no). Cada una de las seis dimensiones mencionadas se puntúa de 0 a 100, correspondiendo 0 cuando se responde negativamente a todos los ítems de una dimensión, y 100 cuando las contestaciones son todas afirmativas. La puntuación final de una dimensión se obtiene dividiendo el número de respuestas positivas por el número total de ítems de esa misma dimensión, multiplicando el resultado por cien. No se calcula en principio una puntuación global, si bien en algunos estudios clínicos se ha

llegado a estimar. En cuanto al tratamiento que reciben las respuestas de la segunda parte del cuestionario, aquellas son analizadas como variables categóricas.

Tabla 2. Dimensiones y puntuaciones resumen calculadas por cada medida*

Dimensiones	SIP	NHP	COOP/WONCA	SF-36/SF-12
Salud general			X	X
Cambio en la salud			X	(X) ^c
Salud física				
Actividades/roles			X	X
Trabajo	X			
Hogar	X			
Ocio	X			
Desplazamientos	X			
Comida	X			
Energía/vitalidad		X		X
Movilidad	X	X	X	X
Dolor		X	(X) ^b	X
Cuidado personal	X			
Sueño/descanso	X	X		
Salud social				
Actividades/roles			X	X
Comunicación	X			
Interacción	X	X		
Salud mental				X
Actividades/roles				X
Atención	X			
Ansiedad/depresión				
Estado emocional	X	X	X	
Puntuaciones resumen				
Salud física	X			X
Salud mental	X			X
Puntuación global	X	(X) ^a		

a. Su cálculo no es recomendado por los autores de la medida, si bien ha sido estimado en algunos estudios.

b. La dimensión/ítem Dolor es opcional.

c. El denominado ítem de transición del SF-36 inquiriere acerca del estado de salud general respecto del año anterior, y no se utiliza para el cálculo de ninguna de las puntuaciones (escalas) de las dimensiones. Este ítem no figura en el cuestionario SF-12.

*: A fin de poder comparar las dimensiones incluidas en cada uno de los instrumentos mostrados en la tabla, los términos empleados para enunciarlas no corresponden exactamente con los utilizados en los cuestionarios reales. Por ejemplo, mientras que en la traducción al español del NHP la dimensión “Movilidad” de la tabla se enuncia como “movilidad física” en las láminas COOP se menciona como “forma física”.

COOP Charts/COOP WONCA Charts

Las láminas o viñetas COOP fueron diseñadas originalmente por un equipo de médicos investigadores en el área de la atención primaria en EE.UU. (*The Dartmouth Primary Care Cooperative Information Project – COOP Project*). La abreviatura COOP proviene de “Cooperative”, denominándose al instrumento *COOP Charts* porque cada una de las opciones de respuesta a los ítems del cuestionario se ilustra mediante una

viñeta. Hay en total 9 dimensiones (Tabla 2), cada una de ellas compuesta por un único ítem. Cada lámina consta de un título en el que se plantea una pregunta referida al estado de salud del paciente en las cuatro últimas semanas, ofreciéndose cinco posibles opciones de respuesta. Como se ha dicho, cada una de estas posibles respuestas se ilustra mediante un dibujo o viñeta que representa el nivel de funcionamiento del sujeto en una escala ordinal tipo Likert (1-5), en la que puntuaciones más altas corresponden a peores niveles de funcionamiento. El cuestionario puede auto-administrarse o ser administrado por un profesional sanitario.

Las láminas COOP-WONCA son el resultado de la revisión efectuada en 1998 por la *World Organization of National Colleges, Academies, and Academia Associations of General Practitioners/Family Physicians* (WONCA) estadounidense. A consecuencia de dicha revisión se acuñó un nuevo instrumento (*COOP-WONCA Charts*) que eliminaba las dimensiones relativas al “apoyo social” y la “calidad de vida” de las láminas COOP originales. Asimismo se dejaba la dimensión “dolor” como opcional. El tiempo de referencia de las preguntas se redujo de cuatro a dos semanas. Por último, los títulos de las láminas, preguntas y viñetas se simplificaron al máximo.

Short Form-36/SF-12

El cuestionario SF-36 fue diseñado por Ware y otros en EE.UU. a principios de los años noventa (Ware y Sherbourne, 1992; Ware et al., 1993) para su utilización en el Estudio de los Resultados Médicos (Medical Outcomes Study). Esto permitió construir un perfil de salud aplicable tanto a población general como a pacientes. El cuestionario abarca 8 dimensiones, cada una de ellas con un número variable de ítems. De manera adicional, el SF-36 incluye un ítem de transición que pregunta por el cambio en el estado de salud general respecto al año anterior. Este ítem no se utiliza para el cálculo de ninguna puntuación, pero proporciona información útil sobre el cambio auto-percibido en la propia salud. Los 35 ítems distribuidos a lo largo de las 8 dimensiones más este último ítem de transición suman los 36 ítems que dan nombre al instrumento. La Tabla 3 muestra la combinación de dimensiones, ítems y niveles de respuesta del SF-36, en las dos versiones existentes del cuestionario.

Las puntuaciones de las 8 dimensiones del SF-36 se ordenan de forma tal que cuanto mayor sea el valor registrado, tanto mejor es el estado de salud correspondiente. Dentro de cada dimensión los ítems son codificados, agregados y transformados en una escala que va de 0 (peor puntuación) a 100 (mejor puntuación). Además, el cuestionario

permite el cálculo de dos puntuaciones resumen, la componente resumen física (PCS) y la componente resumen mental (MCS), mediante la suma ponderada de las puntuaciones de las 8 dimensiones principales. Las ponderaciones empleadas se obtuvieron a partir de estudios de análisis factorial realizados con diferentes muestras de la población general estadounidense (Ware et al., 1994). La principal ventaja que ofrece fundamentar las puntuaciones resumen en algoritmos derivados de una población de referencia, es que los resultados obtenidos son directamente interpretables respecto a la misma. En concreto, tal y como puede apreciarse en la Tabla 3, puntuaciones superiores o inferiores a 50 indican un mejor o peor estado de salud, respectivamente, que la población de referencia.

Tabla 3. Rasgos principales de los cuestionarios SF-36 y SF-12*

Dimensiones	Ítems		Niveles		Puntuación		
	SF-36	SF-12	SF-36v1	SF-36v2	SF-36v1	SF-36v2	SF-12
Función física	10	2	21	21	0-100	50 (10)	50 (10)
Rol físico	4	2	5	17	0-100	50 (10)	50 (10)
Dolor corporal	2	1	11	11	0-100	50 (10)	50 (10)
Salud general	5	1	21	21	0-100	50 (10)	50 (10)
Vitalidad	4	1	21	17	0-100	50 (10)	50 (10)
Función social	2	1	9	9	0-100	50 (10)	50 (10)
Rol emocional	3	2	4	13	0-100	50 (10)	50 (10)
Salud mental	5	2	26	21	0-100	50 (10)	50 (10)
PCS	21	6	--	--	50 (10)	50 (10)	50 (10)
MCS	14	6	--	--	50 (10)	50 (10)	50 (10)

*: Las puntuaciones 0-100 tienen ese rango, interpretándose como el peor o mejor estado de salud, respectivamente. En las puntuaciones 50 (10) la población de referencia tiene una puntuación 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.

Fuente: Adaptado de Vilagut et al. (2005).

A finales de los años noventa se desarrolló la versión 2.0 del cuestionario original SF-36 con el objetivo, entre otros, de mejorar las características métricas de las escalas de rol físico y emocional (Ware et al. 2000). Como resumen Vilagut et al. (2005), los cambios principales introducidos por la versión 2.0 del SF-36 han sido: a) instrucciones y enunciados más sencillos; b) mejoras en la presentación de los ítems y las opciones de respuesta; c) mayor comparabilidad en las diferentes traducciones del cuestionario; d) cinco 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.

Asimismo, esta versión utiliza algoritmos de puntuación basados en las normas poblacionales (de forma análoga a las puntuaciones resumen) para las 8 escalas del SF-36 (véase Tabla 3).

La adaptación transcultural del SF-36 ha sido abordada por investigadores de 15 países diferentes, incluida España, como parte del *Internacional Quality of Life Assessment (IQOLA) Project* (Aaronson et al. 1992, Ware et al. 1995 y 1996a y Gandek y Ware 1998). Las adaptaciones validadas y publicadas por el proyecto IQOLA incluyen el castellano (Alonso et al. 1995), el francés (Perneger et al. 1995), el alemán (Bullinger 1995), el sueco (Sullivan et al. 1995) y el italiano (Apolone et al. 1997), así como diferentes variaciones del inglés para su uso en Canadá, Reino Unido, Australia y Nueva Zelanda (McCallum 1995). Actualmente ya se ha desarrollado la versión 2.0 del cuestionario en castellano, aunque por el momento no hay datos sobre su validación, ni estudios comparativos con la versión original (Vilagut et al., 2005).

El cuestionario SF-12 surgió (Ware et al., 1996) como un intento de acortar el tiempo requerido para la administración del cuestionario SF-36, sin merma significativa de la capacidad explicativa de las puntuaciones resumen. Este instrumento está formado por una selección de 12 ítems del conjunto total que integra el SF-36 (1-2 por dimensión). En este sentido, se ha constatado empíricamente (Vilagut *et al.*, 2005) que las medidas resumen física y mental del SF-12 (PSC-12 y MCS-12 respectivamente) logran predecir con bastante precisión las puntuaciones resumen del SF-36. Hay dos versiones del cuestionario (Ware et al., 2002), siendo sus principales diferencias las siguientes: a) la versión 1.0 sólo permite calcular las puntuaciones resumen PCS-12 y MCS-12, pero no una puntuación por dimensión, cosa que sí ofrece la versión 2.0; b) la versión 2.0 incorpora mejoras en la presentación y en las instrucciones que presuntamente facilitan la auto-administración del cuestionario; y c) las opciones de respuesta dicotómica (sí/no) de la versión 1.0 son reemplazadas en la versión 2.0 por opciones multinivel.

2.2.3. Propiedades empíricas del SF-36 en comparación a otras medidas

Hay varias revisiones bibliográficas que han analizado el grado de utilización de diferentes medidas de CVRS así como sus propiedades psicométricas, dentro (Vilagut et al., 2005) y fuera de España (Brazier et al., 1999; Garratt et al., 2002). La revisión metodológica abordada en este proyecto, por su condición de tal, no puede (ni lo pretende) ofrecer por si misma una panorámica completa de la diversidad de aplicaciones y propiedades del cuestionario SF-36, si bien si que hemos accedido a

varios estudios en los que se compara dicha medida con algunas de las que hemos descrito en el apartado anterior. Dicha comparación se ha realizado comúnmente sobre una base psicométrica. Resumimos a continuación los principales resultados que hemos hallado a este respecto.

Son varios los artículos que se hacen eco de la revisión bibliográfica realizada por Garratt et al. (2002), quienes analizaron 2921 publicaciones en las que se aborda el desarrollo y contraste de medidas de CVRS aplicadas a pacientes. Aunque la mayoría de esas publicaciones informan de la aplicación de alguna medida específica de CVRS (1819), los perfiles de salud fueron la segunda categoría más empleada (690), por delante de las medidas basadas en preferencias (409). Si se atiende a la frecuencia con que se evaluaron las propiedades de una determinada medida, el SF-36 resulta ser la más frecuente (10% del total de publicaciones). Si a esta se le añade el SIP y el NHP, entre las tres concentran el 16% del total de publicaciones.

En cuanto a las propiedades psicométricas habitualmente evaluadas en la literatura, podemos concluir que son las cuatro siguientes: factibilidad, fiabilidad, validez y sensibilidad al cambio. Por factibilidad o viabilidad se hace referencia a la brevedad y claridad del instrumento de medición, de forma que se minimice la carga cognitiva de los encuestados, así como los costes de recolección y procesamiento de los datos. Este rasgo suele evaluarse atendiendo al tiempo promedio requerido para completar el cuestionario, así como al porcentaje promedio de respuestas faltantes (*missing items*). La propiedad de fiabilidad o consistencia nos informa del grado de precisión (ausencia de error) del instrumento de medida. Esta cualidad puede medirse al menos de dos formas complementarias. En primer lugar, puede medirse la consistencia interna u homogeneidad del cuestionario, examinando en qué medida ítems similares dan lugar a respuestas semejantes. Para este fin suele calcularse el estadístico *alfa de Chronbach*, cuyo valor recorre un intervalo 0 (ausencia de consistencia)-1 (consistencia total). La reproducibilidad del cuestionario o fiabilidad *test-retest* indica el grado de consistencia entre mediciones repetidas efectuadas en distintos momentos del tiempo. Para verificar este tipo de fiabilidad se calculan Coeficientes de Correlación Intraclase (CCI), también en un rango 0-1. En cuanto a la validez, ésta puede revestir varias formas, pero como no existe un *gold standard* para medir la salud, la validación psicométrica de un cuestionario de CVRS es necesariamente indirecta. Las dos propiedades que a este respecto se examinan son las siguientes: por un lado, la validez convergente del instrumento, por comparación de sus resultados con los obtenidos con indicadores

clínicos y con otras medidas de salud auto-percibida; y, por otro lado, su validez discriminante, contrastando la capacidad del cuestionario para discriminar entre distintos grupos de pacientes, los cuales se espera que difieran en su estado de salud. De nuevo suelen calcularse coeficientes de correlación para verificar lo primero (validez convergente), mientras que en el segundo caso (validez discriminante) se construyen también curvas ROC (*Receiver Operating Characteristic*). Finalmente, diremos que un cuestionario es sensible al cambio si es capaz de detectar cambios en el estado de salud cuando estos se producen realmente. La magnitud de este cambio suele medirse mediante el cálculo del llamado tamaño del efecto (*size effect*), consistente en la división de la diferencia entre las puntuaciones medias antes y después de la intervención entre la desviación estándar de los cambios en las puntuaciones.

La revisión efectuada permite afirmar que el SF-36 posee buenas propiedades psicométricas, variando lógicamente en función del contexto en el que se haya pretendido validar el cuestionario. En términos generales, podríamos decir que manifiesta una viabilidad media-alta, con un tiempo medio requerido para responder a las preguntas (5-10 minutos) inferior al SIP, y semejante al NHP, aunque superior al empleado con las láminas COOP/WONCA (4-6 minutos). Sin embargo, el porcentaje de respuestas faltantes es, en promedio, superior a instrumentos como el NHP o el COOP/WONCA. Su fiabilidad es, en cambio, bastante elevada, superior por lo general a la registrada por el resto de instrumentos, y por encima de los estándares admitidos (alfa de Chronbach y CCI superiores a 0.7). El SF-36 ofrece una validez convergente razonable, con una capacidad notable para discriminar entre diferentes grupos de pacientes. Por último, el cuestionario SF-36 goza de una gran sensibilidad al cambio (probablemente su mayor virtud), superior en promedio a la de cualquiera de los instrumentos de la Tabla 2. Nuestra revisión nos ha permitido encontrar diversos estudios en los que se ha validado el SF-36 en diferentes grupos de pacientes. Algunos ejemplos son: Hacking et al. (2006) aplican el cuestionario a pacientes que han sufrido embolias, Ünala et al. (2001) a pacientes con afecciones hepáticas, Essink-Bot et al. (1997) lo hacen con enfermos crónicos de migraña, Maori et al. (2001) con pacientes renales dializados, Juniper et al. (2001) en el contexto de los enfermos de asma.

Para España disponemos de dos fuentes de información básicas que permiten valorar psicométricamente la traducción al castellano del SF-36. Vilagut et al. (2005) revisó 17 artículos en los que se habían estudiado alguna de las propiedades que hemos visto

antes con la versión española del SF-36. Sus conclusiones son positivas para el SF-36, habiendo constatado una alta consistencia interna, buena capacidad discriminante entre grupos de pacientes de gravedad dispar, correlación moderada con indicadores clínicos y elevado grado de asociación con otros instrumentos de CVRS. Por otra parte, Badía et al. (1996) desarrollaron un índice sintético para valorar la calidad global de las medidas de CVRS para su uso en España. Dicho indicador (índice GRAQoL) está compuesto por 11 ítems (traducción, validación, sensibilidad, etc.) de respuesta dicotómica (0-2). El valor del índice se calcula mediante la división de la suma de las puntuaciones otorgadas a los diferentes ítems entre la máxima puntuación posible, expresado finalmente como un porcentaje. Un valor del índice GRAQoL inferior al 50% indica un escaso grado de desarrollo del instrumento en cuestión en nuestro país; entre 51% u 70%, un nivel de desarrollo aceptable; y del 71% en adelante, un nivel de desarrollo bueno. De las medidas genéricas que hemos venido considerando hasta el momento (las de la Tabla 2), Badía et al. (1996) puntúan tres: el SIP, el NHP y el SF-36. El valor del índice GRAQoL en cada caso es, respectivamente, 90%, 100% y 77%, con lo cual los tres instrumentos mencionados gozan de un buen nivel de desarrollo en España.

2.2.4. Medidas basadas en preferencias

Como ya se señaló en el apartado 2.2.1., hay medidas genéricas de la CVRS que añaden al componente descriptivo multidimensional que las define, un algoritmo generador de índices cardinales de preferencias, también denominadas utilidades. El objetivo fundamental de este proyecto ha sido precisamente estimar dicho algoritmo para una medida de utilidad multiatributo, el SF-6D, directamente emparentada con el SF-36, ya que sus atributos y niveles se definen a partir de los ítems y dimensiones de este último cuestionario. ¿Por qué son necesarias las medidas basadas en las preferencias? ¿Por qué no nos basta con disponer de sistemas descriptivos como los que ya hemos examinado? La revisión efectuada arroja algunas respuestas sobre estas cuestiones:

- Si se pretende resumir en un solo número el estado de salud de un individuo (y por agregación de la población), resulta evidente que sólo podremos utilizar aquellas medidas de CVRS que incorporen una fórmula de puntuación tal que cualquier combinación de niveles del conjunto proporcione una valoración global. Aunque la mayoría de las medidas no basadas en preferencias ofrecen puntuaciones por dimensiones, e incluso por agrupaciones de dimensiones, casi

ninguna permite obtener un valor resumen global. Como hemos tenido ocasión de comprobar, sólo el SIP ofrece este recurso.

- Para que puedan emplearse las puntuaciones obtenidas de las medidas de CVRS en el cálculo de Años de Vida Ajustados por la Calidad (AVAC)⁵, se requerirá que tales medidas cumplan dos condiciones: a) que puedan combinarse con el tiempo de vida; b) que produzcan valores en una escala intervalo definida por el 1 (salud completa) y el 0 (la muerte). Esto es así porque los AVAC son precisamente combinaciones de cantidad y calidad de vida, donde los pesos, ponderaciones o utilidades de la calidad de vida se miden en una escala de esas características. Esta propiedad es la que permite que pueda compararse entre sí la efectividad de diferentes programas sanitarios. Las escalas psicométricas no ofrecen solución al problema de la combinación de la cantidad y la calidad de vida, ni tampoco producen valores cardinales, sino meramente ordinales.
- Por último, la mayor limitación de las medidas no basadas en preferencias reside precisamente en que no reflejan las preferencias de la sociedad. De hecho, los AVAC son considerados mayoritariamente funciones de utilidad⁶. Por esta razón, desde el punto de vista económico se considera que las reglas de agregación de las puntuaciones que subyacen a las escalas psicométricas son arbitrarias⁷. Por ejemplo, el cuestionario SF-36 supone que cuando un encuestado afirma que el dolor le ha dificultado su trabajo habitual “un poco”, la distancia entre ese nivel de la escala y el nivel “nada”, es el mismo que existe entre los niveles “bastante” y “mucho”. Parece claro que este supuesto dista mucho de ser auto-evidente, ya que dicha igualdad en la ponderación entre niveles adyacentes dependerá de cuál sea la verdadera intensidad de preferencias del sujeto. Su preferencia por pasar de tener “un poco” de dolor a no tener “nada” de dolor podría ser, por ejemplo, mayor (más intensa) que su preferencia por pasar de tener “mucho” dolor a tener “bastante” dolor.

Una vez que hemos mostrado las ventajas que proporcionan las medidas basadas en preferencias, queda por dilucidar la cuestión de cómo pueden ponerse en relación dos instrumentos como el SF-36 y el SF-6D. La conexión o “puente” entre ambas medidas

⁵ Varios trabajos abordados por el equipo investigador analizan los fundamentos teóricos y aplicaciones empíricas de los AVAC (p.ej. Abellán y Pinto, 2000; Pinto y Sánchez, 2003; Abellán et al., 2006).

⁶ Véase, por ejemplo, Pliskin et al. (1980), Bleichrodt y Quiggin (1997), Abellán y Pinto (2000).

⁷ Puede acudirse a Brazier et al. (1999) para consultar una extensa discusión sobre este particular.

se explicará en detalle en el epígrafe 2.2.5; baste por el momento describir con carácter general aquellas medidas en preferencias que hemos identificado en nuestra revisión como más importantes, incluyendo entre ellas al SF-6D. Atenderemos para ello a los dos componentes que caracterizan a una medida de CVRS basada en preferencias, a saber: a) el sistema descriptivo que incorporan; y b) la modelización que se ha empleado para obtener el conjunto de utilidades representativas de los estados de salud del instrumento (la “tarifa” social).

2.2.4.1. El componente descriptivo.

La Tabla 4 muestra el número de atributos o dimensiones de la salud que abarcan los instrumentos que hemos identificado en la revisión, los niveles de gravedad que componen dichas dimensiones, y el número de estados de salud que pueden llegar a describir. De las medidas reunidas en la tabla, sirvan como ejemplo los que probablemente sean los tres instrumentos más empleados en la actualidad (Brazier et al., 1999): *Health Utility Index*, *EQ-5D* y *SF-6D*.

Tabla 4. Componente descriptivo de algunos sistemas de utilidad multiatributo

Sistema descriptivo	Atributos	Niveles por atributo	Estados de salud
QWB ^a	4	3, 5 y 27	1215
Rosser ^b	2	4 y 8	29
15-D ^c	15	5	> 30 millones
HUI3 ^d	8	5 ó 6	972000
EQ-5D ^e	5	3	243
AQoL ^f	15	4	> 1 millón
SF-6D (SF-36) ^g	6	4 a 6	18000
SF-6D (SF-12) ^h	6	3 a 5	7500

a. The Quality of Well-Being Scale (Patrick *et al.*, 1973)

b. Rosser y Kind (1978)

c. Sintonen (1994)

d. Los antecedentes del Health Utility Index Mark 3 (Feeny *et al.*, 1998) son el HUI1 (Torrance et al., 1982) y el HUI2 (Torrance et al., 1996).

e. EuroQol-5D (EQ-5D), Brooks (1996).

f. The Assessment of Quality of Life (Hawthorne *et al.*, 1997).

g. Brazier *et al.* (2002).

h. Brazier y Roberts (2004).

Fuente: Adaptado de Dolan (2002).

Health Utility Index

Torrance et al. (1982) desarrollaron un instrumento conocido como *Health Utility Index Mark 1* (HUI1) para su uso en la evaluación de los resultados de salud producidos por la

asistencia intensiva a neonatos. Dicho sistema posee una estructura multiatributo de cuatro dimensiones y un rango de niveles que va de 4 a 8: funcionamiento físico (6 niveles), función de rol (5), función socio-emocional (4) y problema de salud (8). Con esta estructura pueden llegar a definirse hasta 960 estados de salud.

El HUI2 (Torrance et al., 1996) incrementó la riqueza informativa del HUI1 original, utilizando siete atributos, con un número variable de niveles (3-5): capacidad sensorial (4 niveles), movilidad (5), emoción (5), cognición (4), cuidados personales (4), dolor (5), y fertilidad (3). El sistema fue diseñado específicamente para describir los efectos de largo plazo sobre la salud de los niños de los tratamientos contra el cáncer. El número total de estados de salud que puede llegar a generarse es de 24000.

Finalmente, el HUI3 (Furlong et al., 1998; Feeny et al., 2002) ofrece un sistema descriptivo aún más completo, compuesto por ocho atributos de 5-6 niveles cada uno: visión (6 niveles), oído (6), habla (5), desplazamiento (6), destreza (6), emociones (5), cognición (6), y dolor (5). El poder descriptivo de este instrumento es tal que puede caracterizar hasta 972000 estados de salud diferentes.

EQ-5D

El origen del sistema EQ-5D fue una revisión sistemática de otras medidas genéricas de CVRS que incluyó al Quality of Well-Being, SIP, NHP y al índice de Rosser, efectuada por el Grupo EuroQol (EuroQol Group, 1990). La versión inicial del sistema actual constaba de seis dimensiones, con 2 ó 3 niveles por dimensión: movilidad (3 niveles), cuidado personal (3), actividad principal (2), actividades de ocio (2), dolor (3), y ansiedad/depresión (2).

Con la información suministrada por los primeros estudios (Essink-Bot et al., 1990; Kind, 1990; Brooks et al., 1991; Nord, 1991) que obtuvieron valoraciones de estados generados por el sistema de seis dimensiones, se analizó en qué medida contribuía cada una de estas dimensiones a explicar dichas valoraciones⁸. La conclusión extraída de estos análisis fue que la contribución del atributo ‘actividades de ocio’ era marginal, optándose finalmente por su integración junto con la dimensión ‘actividad principal’ dentro de una única dimensión denominada ‘actividades cotidianas’. Asimismo, todas las dimensiones fueron divididas en tres niveles –ningún problema, algunos problemas,

⁸ Este tipo de análisis suele hacerse especificando un modelo aditivo multiatributo en el que los niveles de las diferentes dimensiones se utilizan como predictores de las valoraciones de los estados de salud (la variable dependiente).

problemas extremos- resultando en una estructura más equilibrada. Nacía así el sistema actual con cinco dimensiones (5D) -movilidad, cuidado personal, actividades cotidianas, dolor/malestar, ansiedad/depresión- y tres niveles de gravedad por dimensión -ningún problema, algunos problemas y problemas extremos. Este sistema puede generar hasta 243 (3^5) estados de salud diferentes, cada uno de los cuales puede ser codificado mediante un número de cinco dígitos. Por ejemplo, el estado 12321 representa una situación caracterizada por la ausencia de problemas de movilidad, la existencia de algunos problemas para ejercer el cuidado personal, muchos problemas para la realización de actividades cotidianas como trabajar o estudiar, el padecimiento de algún dolor o malestar, y la inexistencia de ansiedad o depresión.

SF-6D

La estructura multiatributo del sistema de clasificación SF-6D se describe en detalle en la sección 2.2.5. Baste decir aquí, simplemente a efectos de comparación con el HUI y el EQ-5D, que desde la primera versión de este instrumento (Brazier et al., 1998) el mismo siempre ha caracterizado la salud mediante una combinación de seis dimensiones: funcionamiento físico, limitaciones de rol, funcionamiento social, dolor, salud mental y vitalidad; cada una de las cuales ha adoptado un número variable de niveles. La motivación fundamental de los investigadores que diseñaron y promovieron esta medida fue la de intentar aprovechar la riqueza informativa (la validez discriminante y la sensibilidad al cambio) del instrumento SF-36 (y posteriormente del SF-12). Los análisis de componentes principales realizados por Brazier et al. (1998, 2002) les permitieron reducir el número de dimensiones del SF-36 de 8 a 6, mediante la exclusión del atributo Salud general y la fusión de las dos dimensiones de Limitaciones en el rol (físico y emocional). Los niveles correspondientes a cada una de las seis dimensiones del SF-6D(SF-36)⁹ provienen de una selección de 11 ítems de los 36 que contiene el SF-36.

2.2.4.2. La modelización.

Para conseguir generar una tarifa o panel de utilidades asociadas a los estados de salud que cada sistema de clasificación multiatributo es capaz de caracterizar, se requiere estimar un modelo. Para realizar dicha estimación hay por lo general dos enfoques

⁹ En lo sucesivo se mantendrá la convención de indicar entre paréntesis a continuación de las siglas SF-6D el cuestionario “base” del que procede (en este caso el SF-36, en otras ocasiones el SF-12).

alternativos (Dolan, 2002; Stevens et al., 2007): uno, el denominado enfoque descompuesto o basado en la teoría de la utilidad multiatributo; y otro, el enfoque compuesto o de inferencia estadística.

El primero de los dos enfoques mencionados requiere que los encuestados valoren cada nivel o ítem de un atributo determinado manteniendo constantes los niveles del resto de atributos (por lo general en el mejor nivel). Las funciones de utilidad uni-atributo resultantes se utilizan para generar las utilidades de estados de salud compuestos (o multiatributo) especificando una función de utilidad multiatributo. Este enfoque es el que utilizó para estimar los algoritmos del HUI. Su principal ventaja radica en que minimiza el número de estados de salud que es necesario valorar de manera directa, para luego poder realizar predicciones *out-of-sample*.

El enfoque compuesto, por su parte, requiere que cada encuestado valore un subconjunto de estados de salud multiatributo. El algoritmo se estima mediante la aplicación de técnicas de regresión lineal, por lo general con modelos de efectos aleatorios. Esta aproximación, más preocupada por la validez empírica que por la teórica¹⁰, es la adoptada en la inmensa mayoría de las estimaciones realizadas con el EQ-5D (Dolan, 1997; Badía et al., 2001; Tsuchiya et al., 2002; Lamers et al., 2006) y con el SF-6D (Brazier et al., 1998; 2002; 2004).

Recientemente Stevens et al. (2007) aplican los dos enfoques a un mismo instrumento, el HUI2, concluyendo que la validez predictiva conseguida con el enfoque de inferencia estadística (esto es, la capacidad que tienen las utilidades predichas por los modelos de coincidir con las utilidades valoradas directamente por la muestra) es superior a la registrada mediante el enfoque de la función de utilidad multiatributo.

2.2.5. Rasgos básicos del SF-6D

La clasificación de estados de salud conocida como SF-6D fue derivada inicialmente por Brazier et al. (1998) a partir del SF-36 y, en esta versión preliminar cada una de las seis dimensiones o atributos admitía entre dos y seis niveles. Un estado de salud SF-6D se define seleccionando un nivel para cada una de las dimensiones: “Funcionamiento físico” (FF), “Limitaciones del rol” (LR), “Funcionamiento social” (FS), “Dolor” (DO), “Salud mental” (SM) y “Vitalidad” (VI), tal y como se muestra en la Tabla 5. La combinación de niveles y dimensiones permitía definir 9000 estados de salud.

¹⁰ En el sentido de que la función que se estima estadísticamente es *ad hoc*, ya que no se hace explícita la función de utilidad multiatributo que subyace a la misma.

Brazier et al. (2002) construyeron una versión revisada del SF-6D a partir del SF-36, en la que se mantenían las mismas seis dimensiones del sistema de clasificación primigenio, si bien se modificaron sustancialmente los niveles dentro de cada dimensión, como puede comprobarse observando la Tabla 6. Con esta versión del SF-6D puede definirse un total de 18000 estados de salud. Dado que los propios autores califican la de 1998 como “versión piloto”, nos referiremos a esta versión de 2002 cuando aludamos al SF-6D(SF-36).

Existe una versión diferente del SF-6D (Brazier et al., 2004a), cuya singularidad radica en que está derivada a partir del SF-12 y no del SF-36, y que se distingue de la descrita en el anterior párrafo en dos aspectos, principalmente. En primer lugar, en la dimensión “Funcionamiento físico”, el SF-6D(SF-12) únicamente contempla tres niveles (frente a los seis de la versión SF-36), según la realización de esfuerzos moderados esté nada, algo o muy limitada. Por otra parte, la dimensión “Dolor” del SF-6D(SF-12) omite la situación “No tiene dolor”, que sí estaba presente como primer nivel en el SF-6D(SF-36). De este modo, en el SF-6D(SF-12) ninguno de los seis atributos admite más de cinco niveles, como se muestra en la Tabla 7.

En la Tabla 8 se muestra la correspondencia entre las dos versiones del sistema de clasificación SF-6D y las preguntas e ítems del SF-36 y del SF-12 de los que se derivan. Todos los ítems del SF-12 pueden extraerse del SF-36, pero no resulta posible pasar directamente del SF-6D(SF-36) al SF-6D(SF-12), debido a tres razones (ver Tabla 9):

- En la dimensión “Funcionamiento físico”, el SF-6D(SF-12) utiliza el nivel 3 de la pregunta 3b del SF-36 (pregunta 2a del SF-12), mientras que en el SF-6D(SF-36) no se emplea.
- En la dimensión “Salud mental”, el SF-6D(SF-12) solo utiliza la pregunta 28 del SF-36 (pregunta 11 del SF-12), mientras que en el SF-6D(SF-36) se emplea también la pregunta 24 del SF-36 (que no tiene equivalente en el SF-12).
- En la dimensión “Dolor”, el SF-6D(SF-36) admite un nivel (“No tiene dolor”) del que carece el SF-6D(SF-12).

Tabla 5. El sistema de clasificación de estados de salud SF-6D (versión “piloto” basada en el SF-36). Brazier et al. (1998)

Nivel	Funcionamiento físico	Limitaciones de rol	Funcionamiento social	Dolor	Salud mental	Vitalidad
1	Su salud no le limita para realizar esfuerzos intensos (p.ej. correr, levantar objetos pesados, participar en deportes agotadores).	No tiene problemas con su trabajo u otras actividades cotidianas a causa de su salud física o de sus problemas emocionales.	Su salud física o sus problemas emocionales no interfieren en absoluto con sus actividades sociales normales.	No tiene dolor.	Se siente muy nervioso o desanimado pocas veces o ninguna.	Se siente exhausto o cansado pocas veces o ninguna.
2	Su salud le limita un poco para realizar esfuerzos intensos (p.ej. correr, levantar objetos pesados, participar en deportes agotadores).	Tiene problemas con su trabajo u otras actividades cotidianas a causa de su salud física o de sus problemas emocionales.	Su salud física o sus problemas emocionales interfieren un poco con sus actividades sociales normales.	Tiene dolor muy suave.	Se siente muy nervioso o desanimado algunas veces.	Se siente exhausto o cansado algunas veces.
3	Su salud le limita para subir varios pisos de escaleras o para caminar más de un milla.		Su salud física o sus problemas emocionales interfieren moderadamente con sus actividades sociales normales.	Tiene dolor suave.	Se siente muy nervioso o desanimado y deprimido bastantes veces.	Se siente exhausto o cansado bastantes veces.
4	Su salud le limita para subir un piso de escaleras o para caminar media milla.		Su salud física o sus problemas emocionales interfieren bastante con sus actividades sociales normales.	Tiene dolor moderado.	Se siente muy nervioso o desanimado y deprimido casi siempre.	Se siente exhausto o cansado casi siempre.
5	Su salud le limita para caminar 100 yardas.		Su salud física o sus problemas emocionales interfieren mucho con sus actividades sociales normales.	Tiene dolor intenso.	Se siente muy nervioso o desanimado y deprimido siempre.	Se siente exhausto o cansado siempre.
6	Su salud le limita para bañarse y vestirse por sí mismo.			Tiene dolor muy intenso.		

Tabla 6. El sistema de clasificación de estados de salud SF-6D(SF-36). Brazier et al. (2002)

Nivel	Funcionamiento físico	Limitaciones de rol	Funcionamiento social	Dolor	Salud mental	Vitalidad
1	Su salud no le limita para realizar esfuerzos intensos (p.ej. correr, levantar objetos pesados, participar en deportes agotadores).	No tiene problemas con su trabajo u otras actividades cotidianas a causa de su salud física o de sus problemas emocionales.	Su salud no le dificulta sus actividades sociales (como visitar a amigos o familiares) en ningún momento.	No tiene dolor.	Nunca se siente muy nervioso o desanimado y deprimido.	Tiene mucha energía siempre.
2	Su salud le limita un poco para realizar esfuerzos intensos (p.ej. correr, levantar objetos pesados, participar en deportes agotadores).	Ha dejado de hacer algunas tareas en su trabajo o en sus actividades cotidianas a causa de su salud física.	Su salud le dificulta sus actividades sociales (como visitar a amigos o familiares) solo alguna vez.	Tiene dolor pero no interfiere con su trabajo habitual (fuera de casa o en las tareas del hogar).	Se siente muy nervioso o desanimado y deprimido sólo alguna vez.	Tiene mucha energía casi siempre.
3	Su salud le limita un poco para realizar esfuerzos moderados (p.ej. mover una mesa, pasar la aspiradora o caminar más de 1 hora).	Hace menos de lo que quisiera hacer a causa de sus problemas emocionales.	Su salud le dificulta sus actividades sociales (como visitar a amigos o familiares) algunas veces.	Tiene dolor que interfiere un poco con su trabajo habitual (fuera de casa o en las tareas del hogar).	Se siente muy nervioso o desanimado y deprimido algunas veces.	Tiene mucha energía algunas veces.
4	Su salud le limita mucho para realizar esfuerzos moderados (p.ej. mover una mesa, pasar la aspiradora o caminar más de 1 hora).	Ha dejado de hacer algunas tareas en su trabajo o en sus actividades cotidianas a causa de su salud física y hace menos de lo que quisiera hacer a causa de sus problemas emocionales.	Su salud le dificulta sus actividades sociales (como visitar a amigos o familiares) casi siempre.	Tiene dolor que interfiere moderadamente con su trabajo habitual (fuera de casa o en las tareas del hogar).	Se siente muy nervioso o desanimado y deprimido casi siempre.	Tiene mucha energía solo alguna vez.
5	Su salud le limita un poco para bañarse o vestirse por sí mismo.		Su salud le dificulta sus actividades sociales (como visitar a amigos o familiares) siempre.	Tiene dolor que interfiere bastante con su trabajo habitual (fuera de casa o en las tareas del hogar).		Nunca tiene mucha energía.
6	Su salud le limita mucho para bañarse o vestirse por sí mismo.			Tiene dolor que interfiere mucho con su trabajo habitual (fuera de casa o en las tareas del hogar).		

Tabla 7. El sistema de clasificación de estados de salud SF-6D(SF-12). Brazier et al. (2004a)

Nivel	Funcionamiento físico	Limitaciones de rol	Funcionamiento social	Dolor	Salud mental	Vitalidad
1	Su salud no le limita para realizar esfuerzos moderados.	No tiene problemas con su trabajo u otras actividades cotidianas a causa de su salud física o de sus problemas emocionales.	Su salud no le dificulta sus actividades sociales (como visitar a amigos o familiares) en ningún momento.	No tiene dolor o el dolor no interfiere con su trabajo habitual (fuera de casa o en las tareas del hogar).	Nunca se siente muy nervioso o desanimado y deprimido.	Tiene mucha energía siempre.
2	Su salud le limita un poco para realizar esfuerzos moderados.	Ha dejado de hacer algunas tareas en su trabajo o en sus actividades cotidianas a causa de su salud física.	Su salud le dificulta sus actividades sociales (como visitar a amigos o familiares) solo alguna vez.	Tiene dolor que interfiere un poco con su trabajo habitual (fuera de casa o en las tareas del hogar).	Se siente muy nervioso o desanimado y deprimido solo alguna vez.	Tiene mucha energía casi siempre.
3	Su salud le limita mucho para realizar esfuerzos moderados.	Hace menos de lo que quisiera hacer a causa de sus problemas emocionales.	Su salud le dificulta sus actividades sociales (como visitar a amigos o familiares) algunas veces.	Tiene dolor que interfiere moderadamente con su trabajo habitual (fuera de casa o en las tareas del hogar).	Se siente muy nervioso o desanimado y deprimido algunas veces.	Tiene mucha energía algunas veces.
4		Ha dejado de hacer algunas tareas en su trabajo o en sus actividades cotidianas a causa de su salud física y hace menos de lo que quisiera hacer a causa de sus problemas emocionales.	Su salud le dificulta sus actividades sociales (como visitar a amigos o familiares) casi siempre.	Tiene dolor que interfiere bastante con su trabajo habitual (fuera de casa o en las tareas del hogar).	Se siente muy nervioso o desanimado y deprimido casi siempre.	Tiene mucha energía solo alguna vez.
5			Su salud le dificulta sus actividades sociales (como visitar a amigos o familiares) siempre.	Tiene dolor que interfiere mucho con su trabajo habitual (fuera de casa o en las tareas del hogar).		Nunca tiene mucha energía

Tabla 8. Equivalencias entre las dos versiones del SF-6D y los cuestionarios de los que se derivan*

Escalas	SF-36v2			SF-12 v2			SF-6D(SF-36)	SF-6D(SF-12)	Comunes	SF-6D
	Ítems	Pregunta	Niveles	Ítems	Pregunta	Niveles	Equivalencia	Equivalencia		
Salud general	1	1	5	1	1	5				
	11a	33	5							
	11b	34	5							
	11c	35	5							
	11d	36	5							
Funcionamiento físico	3a	3	3				1			Funcionamiento físico
	3b	4	3	2a	2	3	2	1	X	Funcionamiento físico
	3c	5	3							
	3d	6	3	2b	3	3				
	3e	7	3							
	3f	8	3							
	3g	9	3							
	3h	10	3							
	3j	12	3							
	4a	13	5							
	4b	14	5	3a	4	5				
Limitaciones de rol (físico)	4c	15	5	3b	5	5	4	2	X	Limitaciones de rol
	4d	16	5							
	5a	17	5							
	5b	18	5	4a	6	5	5	3	X	Limitaciones de rol
	5c	19	5	4b	7	5				

Tabla 8 (cont.). Equivalencias entre las dos versiones del SF-6D y los cuestionarios de los que se derivan *

Escala	SF-36v2			SF-12 v2			SF-6D(SF-36)		SF-6D(SF-12)		Comunes	SF-6D
	Ítems	Pregunta	Niveles	Ítems	Pregunta	Niveles	Equivalencia	Equivalencia	Equivalencia	Equivalencia		
Funcionamiento social (extensión)	6	20	5									Dimensiones
Dolor (magnitud)	7	21	6	5	8	5	6	4	X			Dolor
Dolor (interferencia)	8	22	5				7					Dolor
Vitalidad	9a	23	5									
	9e	27	5	6b	9	5	9	5	X			Vitalidad
	9g	29	5									
	9l	31	5									
Salud mental	9b	24	5				8					Salud mental
	9c	25	5									
	9d	26	5	6a	10	5						
	9f	28	5	6c	11	5	10	6	X			Salud mental
	9h	30	5									
Funcionamiento social (frecuencia)	10	32	5	7	12	5	11	7	X			Funcionamiento social

* A partir del SF-36 v2 pueden extraerse las respuestas a los ítems del SF-12 vs. No es posible, sin embargo, pasar del SF-6D(SF-36) al SF-6D(SF-12), pues, en primer lugar, el segundo tiene 3 niveles menos en la dimensión "Funcionamiento físico", uno de los cuales, además, no coincide con los del SF-6D(SF-36). En segundo lugar, el SF-6D(SF-12) tiene un nivel menos en la dimensión "Dolor" (ver Tabla 10).

Tabla 9. Discrepancias entre el SF-6D(SF-36) y el SF-6D(SF-12)*

Funcionamiento físico (FF)		Limitaciones de rol (LR)		Funcionamiento social (FS)		Dolor (DO)		Salud mental (SM)		Vitalidad (VI)	
SF-36	SF-12	SF-36	SF-12	SF-36	SF-12	SF-36	SF-12	SF-36	SF-12	SF-36	SF-12
	1								1 2 3 4		
		1	1	1	1	1		1		1	1
		2	2	2	2	2	1	2		2	2
1		3	3	3	3	3	2	3		3	3
2		4	4	4	4	4	3	4		4	4
3	2			5	5	5	4	5		5	5
4	3						6				
5											
6											

*: Los dígitos que se muestran dentro de la tabla corresponden a los niveles que cada uno de los dos instrumentos posee en cada dimensión.

2.2.6. Modelizaciones de las utilidades SF-6D

Se señalaba más arriba que, a pesar de las virtudes acreditadas del SF-36 y del SF-12 en términos de fiabilidad, validez y sensibilidad al cambio, resultaba imposible utilizar sus puntuaciones como pesos o utilidades de CVRS en el cálculo de los AVAC. El SF-6D (en sus dos versiones), gracias a su mayor simplicidad, ha permitido salvar este obstáculo, posibilitando la generación de tarifas de utilidades basadas en las preferencias de amplias muestras de la población general, susceptibles de ser aplicadas a las respuestas o puntuaciones otorgadas al SF-36/SF-12, mediante el correspondiente algoritmo. Se exponen a continuación las modelizaciones de las utilidades SF-6D existentes hasta la fecha en la literatura: Brazier et al. (1998), Brazier et al. (2002), Brazier et al. (2004a), Kharroubi et al. (2005; 2007a; 2007b) y McCabe et al. (2006).

2.2.6.1. Brazier et al. (1998)

Se trata del primer intento de obtener un índice sintético para el SF-36 basado en preferencias mediante la simplificación o “reestructuración” del SF-36, usando criterios explícitos, para elaborar la clasificación de estados de salud SF-6D. Esta versión SF-6D es la que antes ha sido calificada como “versión piloto” (ver Tabla 5).

El paso inicial en el procedimiento de estimación consistió en la selección de un subconjunto de estados de salud definidos según el SF-6D al objeto de su valoración directa, sobre la que se obtendrían, por extrapolación, las valoraciones de los restantes

estados. La selección de estados se realizó de tal modo que: a) se asegurase que cada nivel de cada dimensión aparecía en al menos uno de los estados; b) en la muestra estuviesen contenidas diferentes combinaciones de niveles y dimensiones, incluidas las que implicasen dificultades físicas, problemas de salud mental y combinaciones de ambos; c) los estados estuviesen presentes en datos reales de ciertos grupos de pacientes y en una muestra de la población general. De este modo se seleccionaron 59 estados.

Un total de 165 personas participó en la valoración de los estados de salud. De ellos, 55 fueron pacientes de consultas externas de reumatología y medicina torácica, así como diabéticos. Los 110 restantes constituyen una muestra no representativa de la población integrada por profesionales sanitarios, gestores y administradores de servicios sanitarios, personal de la Escuela de Medicina de la Universidad de Sheffield y estudiantes de medicina y economía de la salud.

Los sujetos “no pacientes” valoraron 12 estados de salud, mientras que los pacientes evaluaron un total de 8 estados cada uno. Los 59 estados seleccionados se distribuyeron del siguiente modo: todos los sujetos evaluaron el estado de salud perfecta, y todos menos los estudiante, valoraron también dos estados “núcleo” (esenciales); 16 de los estados fueron valorados por un 20% de los entrevistados; los 40 restantes fueron evaluados por un 10% de la muestra.

Las técnicas de valoración de los estados de salud empleadas fueron dos: la escala visual analógica (VAS) y la “lotería estándar” (SG). En el primer caso, se utilizó la versión desarrollada por el grupo Euroqol, y los valores asignados en la escala a cada estado de salud Q se normalizaron en una escala donde 1 se identifica con la salud perfecta y 0 con la muerte. Las preguntas de la lotería estándar respondieron al siguiente “framing”:

(Q, 10 años) vs. (SN, 10años; p ; Muerte inmediata)

siendo p el valor de la probabilidad que consigue la indiferencia entre ambos resultados.

La modelización se planteó inicialmente mediante un modelo de regresión simple en el que las puntuaciones del VAS ajustadas y las utilidades SG eran las variables dependiente, mientras que los niveles de las dimensiones de cada estado se definían como “dummies”:

$$y_i = \alpha + \sum_{k=1}^n \sum_{j=1}^6 \beta_{kj} X_{kj} + e_i \quad (1)$$

El modelo tiene un término constante y un conjunto de coeficientes para el nivel k (de 1 hasta n) en cada dimensión j (de 1 hasta 6) de la clasificación. Con el fin de tener en cuenta el hecho de que los sujetos no valoraron todos los estados y que cada uno de éstos fue evaluado por un subconjunto de individuos diferente, los autores modificaron el modelo introduciendo efectos fijos:

$$y_i = \alpha_0 + \alpha_r + \sum_{k=1}^n \sum_{j=1}^6 \beta_{kj} X_{kj} + e_i \quad (2)$$

Donde α_0 es la constante global y α_r representa la constante específica de cada sujeto. La posible existencia de interacciones entre dimensiones se discute en el artículo sin que se presenten resultados de estimaciones que incluyan tales interacciones.

Tabla 10. Estimaciones de los modelos “consistentes” de efectos fijos para el SF-6D (“versión piloto”) (Brazier et al. 1998)

Modelo VAS			Modelo SG		
Niveles/ dimensiones	Coefficiente	t	Niveles/ dimensiones	Coefficiente	t
Constante	0.425	-25.8***	Constante	0.150	-11.3***
FF2	-0.785	-5.8***	FF2	-0.012	-0.9
FF3	-0.120	-7.2***	FF3	-0.028	-2.1*
FF4	-0.172	-8.8***	FF45	-0.064	-4.6***
FF5	-0.232	-10.5***			
FF6	-0.279	-12.4***	FF6	-0.098	-5.4***
LR2	-0.088	-5.1***	LR2	-0.025	-1.8
FS3	-0.033	-2.0***	FS3	-0.025	-1.7
FS45	-0.094	-5.0***	FS45	-0.033	-2.5*
DO23	-0.017	-1.1	DO23	-0.022	-1.6
DO4	-0.082	-3.0***	DO4	-0.024	-1.8
DO56	-0.120	-4.4***	DO5	-0.129	-6.0***
			DO6	-0.163	-5.7***
SM2	-0.070	-4.6***	SM2	-0.021	-1.9
SM3	-0.115	-5.7***	SM34	-0.032	-2.3*
SM45	-0.125	-5.8***			
SM5			SM5	-0.193	-8.0***
VI2	-0.036	-2.2*			
VI34	-0.074	-3.5***	VI345	-0.020	-1.8
VI5	-0.100	-3.4***			
R ² aj.	0.492		R ² aj.	0.495	

* Significativo al 10%. ** Significativo al 5%. *** Significativo al 1%.

Para predecir los valores de los estados de salud, se debe computar la constante teniendo en cuenta el efecto sujeto medio. Esto sitúa el término constante en 0.92 para el modelo VAS y en 0.99 para el modelo SG.

Tras estimar los modelos de efectos fijos a partir de los valores del VAS y del SG, respectivamente, los autores reestiman los modelos combinando los pares de coeficientes de niveles adyacentes que ponen de manifiesto inconsistencias, esto es, aquellos casos en los que el coeficiente asociado a un nivel en una dimensión es menor en valor absoluto (es decir, resta menos utilidad) que el coeficiente del nivel inmediato inferior. Estos modelos “consistentes” conservan casi intacto el poder explicativo de los originales y sus resultados se resumen en la Tabla 10.

2.2.6.2. Brazier et al. (2002)

Para esta segunda modelización, los autores elaboraron, en primer lugar, una versión revisada del sistema de clasificación SF-6D a partir del SF-36 (ver Tabla 6). Otra diferencia respecto del estudio pionero tiene que ver con la selección de los sujetos que habrían de valorar los estados seleccionados. En este caso se utilizó una muestra representativa de la población general, extraída mediante procedimientos aleatorios, toda vez que el propósito final del proyecto era el de informar la asignación de recursos públicos. La muestra definitiva de sujetos que respondieron al cuestionario estuvo compuesta por 836 individuos representativos por grupo de edad, nivel educativo y clase social de la población británica.

Por lo que respecta a la selección de estados de salud que se valoraron directamente por parte de los sujetos, se identificó el conjunto mínimo de estados mediante un diseño ortogonal, que generó los 49 estados (de un total de 18000 posibles) necesarios para estimar un modelo aditivo. Teniendo en cuenta que sería necesario ensayar especificaciones más complejas que permitiesen interacciones entre dimensiones, se consideró conveniente ampliar la selección de estados. Otra razón para ello fue la de proporcionar un espectro mayor sobre el que testar la validez predictiva de los modelos. Los estados de salud se clasificaron como “leves”, “moderados” o graves” y se aplicó un método de muestreo estratificado para seleccionar 200 estados adicionales que, unidos a los generados por el procedimiento de ortogonalización, sumaron los 249 que finalmente serían valorados por los sujetos que componían la muestra. Cada sujeto hubo de valorar cinco estados de salud y, adicionalmente, el “peor estado” que se puede describir con el sistema de clasificación (645655). Los cinco estados se distribuyeron entre los entrevistados de modo que en los subconjuntos no predominasen los estados “buenos” o los “malos” y de manera que cada uno de los 249 fuese evaluado por un mismo número de sujetos.

Para la valoración directa de los 249 estados de salud se utilizó el método de la lotería estándar (SG). Los sujetos debían elegir entre la certeza de vivir en el estado de salud intermedio que se pretendía valorar y una “lotería” compuesta por dos posibles resultados: el mejor y el peor estado de salud definidos por el SF-6D. La distribución de probabilidad de la lotería que alcanzaba la indiferencia entre ambas alternativas se buscó mediante un procedimiento de “ping pong”. Para transformar los valores en utilidades (en una escala donde la muerte toma el valor 0 y la salud perfecta el valor 1), fue necesario valorar a continuación el peor estado respecto a la muerte utilizando el mismo método de la lotería estándar (con un *framing* diferente según dicho estado fuese considerado por el sujeto como mejor o peor que la muerte inmediata).

La utilización del peor estado como referencia, en lugar de la muerte, se fundamentó, en primer lugar, en la insensibilidad que el método de valoración mostró en la parte superior de la escala en tests previos (la versión del método empleada permitía considerar probabilidades de hasta 0.95 y muchos sujetos sólo estaban dispuestos a elegir la alternativa incierta para probabilidades mayores); el proceso en dos etapas permitía que los sujetos revelasen valores superiores a 0.975. En segundo lugar, el “encadenamiento” evitaba tener que incorporar en cada pregunta la opción de que el estado se considerase peor que la muerte, pues esto se hacía sólo una vez al final, con ocasión de la valoración del peor estado de salud.

En consecuencia, siendo SG el valor asignado al estado de salud intermedio en la lotería estándar frente a salud normal y al “peor” estado, y siendo P el valor que para éste se obtiene al evaluarlo con referencia a salud normal y muerte, los valores ajustados ($SGADJ$) que se utilizaron en la modelización se obtuvieron como:

$$SGADJ = SG + (1-SG) P \quad (3)$$

De los 836 sujetos entrevistados, 225 fueron excluidos del análisis. No completar la segunda etapa de valoración (no valorar el “peor” estado SF-6D) fue la causa de exclusión de 130 individuos. Otros 9 fueron excluidos por haber omitido la valoración de dos o más estados de salud. Por último, se identificaron 86 sujetos cuyas valoraciones fueron coincidentes para los cinco estados, razón por la que se decidió, así mismo, su exclusión. Se mantuvo en la muestra, por el contrario, a los sujetos que incurrieron en inconsistencias lógicas en sus valoraciones, esto es, quienes violaban en sus respuestas las propiedades ordinales del SF-6D, con el fin de maximizar el número de observaciones. En consecuencia, 611 individuos conformaron la base de datos

definitiva, lo que se tradujo en un total 3518 valoraciones SG para 249 estados de salud, con una media de 15 valoraciones por estado.¹¹

La elección del modelo para estimar la tarifa tuvo en cuenta las características de los datos obtenidos de las valoraciones directas. La distribución de estas valoraciones resultó ser asimétrica y bimodal, lo que pone en duda, entre otras cosas, la idoneidad de utilizar la media y no la mediana como medida de tendencia central. Por otro lado, el diseño del estudio dio lugar a que los datos presentasen variabilidad entre sujetos e intra-sujetos (entre estados). Los sujetos, además, no valoraron el mismo conjunto de estados, con lo que las diferencias entre las valoraciones de los estados pudieran deberse, en parte, a diferencias en las preferencias de los sujetos, y no solo a diferencias en los atributos de los estados. Aunque la identificación de los efectos-sujeto requiere la estimación con datos individuales (cada valoración se considera una observación separada), los autores ponen en duda la superioridad de esta modelización frente a la basada en medias por estado, por lo que decidieron estimar modelos a escala individual y a escala agregada.

El modelo general se formuló en los siguientes términos:

$$y_{ij} = g(\beta' \mathbf{x}_{ij} + \theta' \mathbf{r}_{ij} + \delta' \mathbf{z}_j) + e_{ij} \quad (4)$$

donde $i = 1, 2, \dots, n$ representa los valores individuales de los estados de salud y $j = 1, 2, \dots, m$ representa a los encuestados. La variable dependiente y_{ij} el valor SG ajustado (SGADJ) para el estado de salud i valorado por el individuo j . El término \mathbf{x} es un vector de *dummies* explicativas ($x_{\delta\lambda}$) para cada nivel λ de la dimensión δ del SF-6D, que toman valores 1 o 0 en función de que, para ese estado, la dimensión δ se encuentre o no, respectivamente, en el nivel λ . Las posibles interacciones entre niveles de diferentes atributos se recogen en el vector \mathbf{r} , mientras que \mathbf{z} es un vector de características personales que pueden afectar al valor que un individuo atribuye a un estado de salud, como el sexo, la edad o el nivel educativo. Por último, g es la función que especifica la forma funcional apropiada y e_{ij} es un término de error cuya estructura y propiedades dependen de los supuestos subyacentes a cada modelo concreto. El modelo aditivo propuesto no impone restricciones sobre la relación entre niveles.

¹¹ La existencia de *missing values* explica la diferencia entre el número de observaciones (3518) y el resultado de multiplicar 611 sujetos por 6 estados des salud (3666).

El punto de partida fue una estimación por MCO del modelo uno suponiendo que g es una función lineal. Para tener en cuenta el hecho de que los datos presentan variabilidad intra-sujetos y entre sujetos, se formuló un modelo de efectos aleatorios¹², en el que los errores del modelo 4 se subdividen del modo siguiente:

$$\varepsilon_{ij} = u_j + e_{ij} \quad (5)$$

siendo u_j la variación específica por sujeto, que se supone distribuida aleatoriamente entre los entrevistados, e_{ij} un término de error para la valoración del estado i del individuo j , que se supone distribuida aleatoriamente a lo largo de las observaciones. La estimación de este modelo se realiza mediante mínimos cuadrados generalizados o máxima verosimilitud.¹³

En relación con los efectos “cruzados” o interacciones entre niveles y dimensiones, se renunció a analizar las interacciones de primer orden por el riesgo de identificar efectos espurios y por la posibilidad de que el hallazgo de efectos cruzados significativos desplazase los efectos principales debido a la colinealidad entre ambos. Por esta razón se buscaron formas alternativas de tener en cuenta interacciones, como la definición de dummies de niveles extremos: número de veces que un estado de salud contiene dimensiones en su nivel mínimo o máximo; presencia de un nivel mínimo/máximo en cualquier dimensión de un estado; presencia del nivel máximo en n dimensiones de un estado. Los resultados de algunas de las estimaciones se muestran en la Tabla 11.

En los modelos de efectos principales, el poder explicativo es mucho mayor para los modelos de medias y medianas que para los modelos individuales (MCO y efectos aleatorios). Existen algunas similitudes entre los cuatro modelos que se presentan, en particular la consistente en que los efectos más importantes se localizan entre los niveles más altos (mayor gravedad) de cada dimensión. La mayor parte de estos efectos son robustos respecto de la especificación del modelo. Por lo que respecta a su capacidad predictiva, los modelos a escala individual y el modelo de medias tienen un error absoluto medio (EAM) similar, mientras que el modelo de medianas parece ser el peor en este aspecto. El principal problema es la existencia de autocorrelación en los errores de predicción en todos los modelos. Se pone de manifiesto una tendencia a predecir por

¹² Los autores contemplan la alternativa de un modelo de efectos fijos, pero consideran que sus datos se ajustan mejor a los supuestos del modelo de efectos aleatorios, lo que, en todo caso, confirmarán con el test de Hausman.

¹³ También consideran la posibilidad de hacer estimaciones mediante MCG iterativos con modelos multi-nivel desarrollados específicamente para tratar con estructuras de datos jerárquicas.

exceso los valores de estados de salud graves y por defecto los de estados de salud leves.

Tabla 11. Estimaciones de los modelos de medias para el SF-6D(SF-36) (Brazier et al., 2002; Brazier et al., 2004a)

	Modelos de efectos principales		Modelos con interacciones		Modelo "consistente"	
	(1)	(2)	(3)	(4)		(5)
<i>Constante</i>	0.827	1.000	0.788	1.000	<i>Constante</i>	1.000
FF2	-0.014	-0.060	-0.015	-0.053	FF23	-0.035
FF3	0.008	-0.020	0.011	-0.011		
FF4	-0.027	-0.060	-0.018	-0.040	FF4	-0.044
FF5	-0.043	-0.063	-0.034	-0.054	FF5	-0.056
FF6	-0.096	-0.131	-0.084	-0.111	FF6	-0.117
LR2	-0.019	-0.057	-0.021	-0.053	LR234	-0.053
LR3	-0.043	-0.068	-0.030	-0.055		
LR4	-0.036	-0.066	-0.024	-0.050		
FS2	-0.027	-0.071	-0.023	-0.055	FS2	-0.057
FS3	-0.049	-0.084	-0.040	-0.067	FS3	-0.059
FS4	-0.057	-0.093	-0.042	-0.070	FS4	-0.072
FS5	-0.073	-0.105	-0.058	-0.087	FS5	-0.087
DO2	0.008	-0.048	0.005	-0.047	DO23	-0.042
DO3	-0.001	-0.034	0.004	-0.025		
DO4	-0.032	-0.070	-0.025	-0.056	DO4	-0.065
DO5	-0.062	-0.107	-0.049	-0.091	DO5	-0.102
DO6	-0.149	-0.181	-0.136	-0.167	DO6	-0.171
SM2	-0.026	-0.057	-0.030	-0.049	SM23	-0.042
SM3	-0.022	-0.051	-0.019	-0.048		
SM4	-0.095	-0.121	-0.089	-0.109	SM4	-0.100
SM5	-0.114	-0.140	-0.109	-0.128	SM5	-0.118
VI2	-0.044	-0.094	-0.044	-0.086	VI234	-0.071
VI3	-0.037	-0.069	-0.031	-0.061		
VI4	-0.029	-0.069	-0.019	-0.054		
VI5	-0.076	-0.106	-0.064	-0.091	VI5	-0.092
<i>MOST</i>			-0.041	-0.070	<i>MOST</i>	-0.061
<i>LEAST</i>			0.048			
R ² aj.	0.583	0.508	0.591	0.526	R ² aj.	0.472
Inconsist.	2	5	1	5		
EAM	0.071	0.074	0.070	0.073	EAM	0.074

Los coeficientes en negrita son significativos al 10%.

Los modelos (2), (4) y (5) se estiman forzando a que la constante sea igual a la unidad. El modelo (5) es una reestimación del (4) en la que se han eliminado las variables con coeficientes no significativos y se han agrupado aquellas cuyos coeficientes daban lugar a inconsistencias.

Las estimaciones en las que se fuerza a que la constante sea igual a la unidad aumentan el número de coeficientes estadísticamente significativos, incrementan ligeramente el número de inconsistencias y, pese a que se detecta un mínimo aumento en el tamaño del

error, se reduce la autocorrelación en los errores (se elimina la tendencia a infra-valorar en las predicciones los estados leves pero se mantiene la sobre-valoración en los estados graves).

De los modelos que incorporan algunos de los efectos cruzados (interacciones) se concluye la existencia de un efecto significativo y negativo a resultas de la presencia de cualquier dimensión en su máximo (peor) nivel, lo que se representa mediante la variable *MOST*. La conclusión es justo la opuesta (aunque este efecto no resulta ser significativo en los modelos con constante unitaria) cuando se trata de la presencia de un nivel mínimo en cualquier dimensión (variable *LEAST*). Los coeficientes de las variables de efectos principales se reducen ligeramente pero son robustos a la inclusión de estas interacciones. No obstante, la mejora en términos de capacidad predictiva de los modelos es mínima.

El poder explicativo de los modelos SF-6D a escala individual es mucho menor que el de los modelos EQ-5D, lo cual no debe sorprender dado que el SF-6D describe un número de estados casi 75 veces mayor que el EQ-5D. Lo que resulta más relevante para el ACU es la capacidad del modelo para predecir valores medios de estados de salud y, en este sentido, el modelo de medias preferido en el estudio de Brazier et al. (2002) tiene un R^2 ajustado de 0.58. Por lo que respecta a las inconsistencias en los resultados de las estimaciones, estas son pocas en términos relativos (comparando los pares de coeficientes inconsistentes con los coeficientes que son consistentes). Algunas de las que aparecen en más de uno de los modelos seleccionados se pueden explicar de manera convincente, por lo que los autores no las consideran un problema importante para el funcionamiento del modelo. Sí se estima preocupante la existencia de errores de predicción sistemáticos.

Por último, los autores identifican como modelo preferido el modelo MCO de medias por estado¹⁴, en el que la constante es igual a la unidad y que incluye, junto a los efectos principales, una única variable representativa de interacciones, la variable *MOST*.

¹⁴ Si bien el modelo de efectos aleatorios es mejor que el MCO a escala individual, los autores no creen que tenga claras ventajas sobre los modelos de medias. Estos, además, se comportan algo mejor en los diferentes tests de ajuste. Dado que el objetivo es predecir valores medios de los estados de salud no hay razón alguna para optar por modelos de datos individuales, de ahí que se recomiende el modelo de medias.

2.2.6.3. Brazier et al. (2004a)

En este caso, el equipo de investigadores abordó la modelización de la tarifa del SF-6D(SF-12), esto es, la versión del sistema descriptivo de seis dimensiones que se deriva del SF-12 (ver Tabla 7). No obstante, la base de datos fue la misma que se utilizó en Brazier et al. (2002), con la salvedad de que, debido a las diferencias en la configuración de las dimensiones “Funcionamiento físico” (FF) y “Dolor” (DO) los 249 estados valorados en el estudio de 2002 se reducen a 241 en esta nueva modelización.

Tabla 12. Estimaciones de los modelos para el SF-6D(SF-12) (Brazier et al., 2004a)

Modelo completo		Modelo consistente	
FF2	0.012		
FF3	-0.040	FF3	-0.045
LR2	-0.068		
LR3	-0.061	LR234	-0.063
LR4	-0.054		
FS2	-0.059	FS2	-0.063
FS3	-0.069	FS3	-0.066
FS4	-0.078	FS4	-0.081
FS5	-0.093	FS5	-0.093
DO2	-0.004		
DO3	-0.039	DO3	-0.042
DO4	-0.076	DO4	-0.077
DO5	-0.140	DO5	-0.137
SM2	-0.063		
SM3	-0.055	SM23	-0.059
SM4	-0.119	SM4	-0.113
SM5	-0.140	SM5	-0.134
VI2	-0.097		
VI3	-0.065	VI234	-0.078
VI4	-0.059		
VI5	-0.103	VI5	-0.106
<i>MOST</i>	-0.085	<i>MOST</i>	-0.077
R ² aj.	0.427	R ² aj.	0.433
Inconsist.	5		
EAM	0.079	EAM	0.079

Los coeficientes en negrita son significativos al 10%.

Ambos modelos se estiman forzando a que la constante sea igual a la unidad. El modelo “consistente” es una reestimación del modelo “completo” en la que se han eliminado las variables con coeficientes no significativos y se han agrupado aquellas cuyos coeficientes daban lugar a inconsistencias.

La estimación de la “tarifa” para el SF-6D(SF-12) se ajustó al modelo “preferido” de entre los estimados en Brazier et al. (2002) para el SF-6D(SF-36), es decir, el modelo MCO de medias con constante unitaria que incluye efectos principales y la variable

MOST. Adicionalmente, se reestimó un modelo “consistente” –y lo mismo se hizo para el SF-6D(SF-36)–, en el que se eliminaron las variables cuyos coeficientes no resultaron estadísticamente significativos en el modelo completo, al tiempo que se agregaron niveles de algunas dimensiones cuando existían inconsistencias (p.ej. como el coeficiente de SM2 era mayor que el de SM3, estas dos *dummies* se sustituyeron por una nueva, SM23).

El poder explicativo de los modelos SF-6D(SF-12) resultó ser similar al de los estimados en 2002 con el SF-6D(SF-36), siendo en general los coeficientes mayores en valor absoluto en el primer caso (con la excepción de las dimensiones funcionamiento físico y dolor). En ambos modelos el dolor parece ser la dimensión más importante para determinar los valores de los estados de salud, seguida de la salud mental. Los errores absolutos medios en las predicciones son pequeños; además, los errores de predicción son insesgados y se distribuyen normalmente. Sin embargo, se confirma el patrón sistemático en los errores que se detectó en Brazier et al. (2002): la tendencia a sobrevalorar los estados graves (peores) en las predicciones.

2.2.6.4. Kharroubi et al. (2005, 2007a, 2007b)

A partir de la misma base de datos utilizada en los dos estudios anteriores (Brazier et al., 2002), Kharroubi et al. (2005) abordan una estimación bayesiana no paramétrica. El modelo bayesiano tiene en cuenta varios aspectos de los datos que sirven de base a la estimación. En particular, este modelo representa de un modo más preciso las siguientes características de los datos:

- (a) la naturaleza de los efectos-sujeto,
- (b) la existencia de medidas repetidas para cada sujeto,
- (c) la distribución asimétrica de las valoraciones individuales para cada estado y
- (d) la relación no paramétrica “suavizada” entre el estado de salud y la utilidad.

El modelo también permite que las covarianzas a nivel de sujeto se integren de manera natural como predictores de los efectos-sujeto. Los autores presentan un sencillo algoritmo computacional “Monte Carlo – cadena de Markov” para computar inferencias posteriores para la media (o mediana) poblacional de la utilidad.

El método se ilustra reanalizando los datos del estudio de valoración del SF-6D (Brazier et al., 2002). Las estimaciones generadas se ajustan a los datos de manera

plausible, repitiendo la monotonía inherente a la medida de utilidad subyacente. El ajuste de modelo se confirma mediante un examen cuidadoso que incluye la predicción de datos no utilizados en la estimación.

Los autores reconocen que la mejora en error medio cuadrático es modesta, pero señalan como otras importantes ventajas el hecho de que su modelo produce:

- (a) predicciones que se comportan mejor debido a que respetan la monotonía subyacente al sistema descriptivo,
- (b) una función de utilidad que captura mejor las utilidades bajas para los estados de salud más graves y
- (c) varianzas “posteriores” que son validadas correctamente frente a las observaciones no utilizadas-

En Kharroubi et al. (2007a) se señalan como principales ventajas teóricas del modelo no paramétrico frente al modelo de efectos aleatorios paramétrico convencional las siguientes:

1. Flexibilidad: No se asume una forma funcional particular para la función de preferencia sobre estados de salud, no siendo necesario hacer supuestos acerca de las interacciones entre niveles y dimensiones.
2. Monotonía. Aunque el modelo no paramétrico permite no-monotonía, en la práctica resulta improbable. De hecho, el grado de no-monotonía es sustancialmente menor en el modelo no paramétrico que en el paramétrico.
3. Valor de la “salud normal”. La respuesta del modelo paramétrico a la conveniencia de asignar valor unitario al estado de “salud perfecta” (forzar la constante a la unidad), da lugar a una pérdida en términos de bondad de ajuste, lo que podría estar detrás de la tendencia del modelo a sobre-predicir los valores de los estados de salud leves. En el modelo no paramétrico, el efecto-sujeto multiplicativo implica que cada encuestado otorga necesariamente el valor 1 a la salud perfecta.
4. Efectos-sujeto. En el modelo no paramétrico, la desviación de la función de preferencia de un individuo respecto del valor poblacional aumenta conforme el estado de salud empeora. Esto supone una mayor variabilidad en los valores asociados a estados de salud graves que en los correspondientes a estados leves,

lo que está en consonancia con lo observado. La aditividad de los efectos sujeto en el modelo paramétrico implica igual variabilidad entre individuos para estados leves y graves.

5. Asimetría. En el SF-6D y en otros sistemas descriptivos se observa asimetría en las valoraciones de los sujetos. Las desviaciones respecto de la media son mayores hacia las utilidades bajas (incluso negativas) que hacia las altas. El uso de la exponencial introduce dicha asimetría en el modelo no paramétrico.

Estas ventajas teóricas se reflejan en las mejoras en términos de capacidad predictiva que se obtienen con el enfoque no paramétrico. En el “debe” de los modelos no paramétricos se ha de anotar, por el contrario, la complejidad computacional que suponen, siendo necesario recurrir a software especializado para estimar la función de preferencia. Además, dicha función no puede definirse como una sencilla tabla de coeficientes, ya que de hecho hay un parámetro para cada uno de los 18000 estados de salud, lo que supone una pérdida de transparencia.

Por último, los autores destacan dos potenciales ventajas adicionales del modelo no paramétrico. La primera consiste en que proporciona un método muy flexible para explorar el impacto de las covarianzas, cuya modelización se aborda en Kharroubi et al. (2007b). Esto tiene un interés intrínseco y puede, además, ser utilizado para estimar tarifas para sub-grupos de población distintos. En segundo lugar, el enfoque bayesiano ofrece la posibilidad de utilizar los resultados de un país para diseñar el modelo en otro país, utilizando aquéllos como *a priori* informativos.

2.2.6.5. McCabe et al. (2006)

McCabe et al. (2006) estiman un modelo de regresión logística utilizando datos de preferencias ordinales. El modelo estima, además, un parámetro para el estado “muerte”, evitando así la necesidad de recurrir a datos externos de utilidad de estados de salud para reescalar los coeficientes, como hace Salomon (2003) en su modelización a partir de datos ordinales del MVH¹⁵. Los datos utilizados por McCabe et al. (2006) son datos ordinales procedentes de sendas encuestas de valoración de estados de salud, una de las cuales es la de Brazier et al., (2002) para el SF-6D.¹⁶

¹⁵ *Measurement and valuation of health study*. Véase Dolan (1997).

¹⁶ La otra es la encuesta del Reino Unido para el HUI2 (McCabe et al., 2005).

Para modelizar los valores de los estados de salud a partir de datos de preferencias ordinales, los autores utilizaron la regresión logística condicionada sugerida por McFadden (1974). Supusieron que el ejercicio de ordenación era equivalente a la realización de una serie de selecciones individuales a partir de conjuntos de estados cada vez más pequeños. Es decir, al ordenar 10 estados de salud se supone que el sujeto primero elige el preferido entre los diez, luego el preferido de entre los nueve restantes y así sucesivamente. Adicionalmente se ha de asumir la hipótesis de independencia de alternativas irrelevantes (el orden entre dos estados no se ve afectado por el resto de estados que se ordenan en el mismo ejercicio).

El modelo de regresión logística condicionada supone que el encuestado i tiene un valor de utilidad latente para el estado j , U_{ij} , y que elegirá el estado j antes que el estado k si $U_{ij} > U_{ik}$. La función de utilidad cardinal de cada individuo para el estado j es:

$$U_{ij} = \mu_j + \varepsilon_{ij} \quad (6)$$

donde μ_j representa los gustos de la población y ε_{ij} los gustos del individuo j . El valor esperado de cada utilidad no observada se supone una función lineal de los niveles y dimensiones de cada conjunto de datos. La especificación general del modelo es, por tanto, la siguiente:

$$U_{ij} = \beta x_{ij} + \theta D + \varepsilon_{ij} \quad (7)$$

donde x es un vector de *dummies* explicativas ($x_{\lambda\delta}$) para cada nivel λ de la dimensión δ del SF-6D. Para un determinado estado de salud, la variable $x_{\lambda\delta}$ toma valor 1 si para ese estado la dimensión δ está en el nivel λ , y toma valor cero en caso contrario. Por su parte, D es una *dummy* que toma valor 1 para el estado “muerte” y cero para el resto de estados. El valor del estado “salud perfecta” se asume igual a la unidad, de manera que el valor de cualquier otro se calcula como 1 menos la suma de los coeficientes para cada variable y nivel del estado en cuestión.

Dado que la variable latente no se estima en la escala 0-1 (muerte-salud perfecta) necesaria para calcular AVAC, los autores reescalaron los coeficientes usando la siguiente fórmula:

$$\beta_{r\lambda\delta} = \beta_{\lambda\delta} / \theta \quad (8)$$

siendo $\beta_{r\lambda\delta}$ los coeficientes reescalados y θ el coeficiente del estado “muerte”. De este modo la muerte se ancla en el valor cero y la salud perfecta en 1, manteniéndose la posibilidad de que un estado de salud tenga un valor negativo (peor que la muerte).

Tabla 13. Estimaciones de los modelos para el SF-6D(SF-36) con datos ordinales (McCabe et al., 2006)

	Coeficientes del modelo de rangos		Coeficientes del modelo SG (Brazier et al.,2002) ^(a)
	Sin reescalar	Reescalados	
FF2	-0.3636	-0.5660	-0.060
FF3	-0.4313	-0.0671	-0.020
FF4	-0.9856	-0.1534	-0.060
FF5	-0.6340	-0.0987	-0.063
FF6	-1.4475	-0.2253	-0.131
LR2	-0.3211	-0.0500	-0.057
LR3	-0.4069	-0.0633	-0.068
LR4	-0.4053	-0.0631	-0.066
FS2	-0.3627	-0.0565	-0.071
FS3	-0.4203	-0.0654	-0.084
FS4	-0.5737	-0.0893	-0.093
FS5	-0.8055	-0.1254	-0.105
DO2	-0.3772	-0.0587	-0.048
DO3	-0.3635	-0.0566	-0.034
DO4	-0.6520	-0.1015	-0.070
DO5	-0.8187	-0.1275	-0.107
DO6	-1.1912	-0.1854	-0.181
SM2	-0.2157	-0.0336	-0.057
SM3	-0.3371	-0.0525	-0.051
SM4	-0.7016	-0.1092	-0.121
SM5	-0.8993	-0.1400	-0.140
VI2	-0.1740	-0.0271	-0.094
VI3	-0.2140	-0.0333	-0.069
VI4	-0.3226	-0.0502	-0.069
VI5	-0.5267	-0.0820	-0.106
Muerte	-6.4240	-1.0000	
EAM		0.088	0.074
EMC		0.110	0.098
Inconsist.		3	5

(a) Modelo de efectos principales estimado a partir de medias forzando a que la constante sea igual a la unidad (columna 2 de la tabla 11).

Los coeficientes en negrita son significativos al 10%.

Los resultados de la estimación se muestran en la Tabla 13, junto con los del modelo de efectos principales basado en medias de Brazier et al. (2002), a efectos comparativos. Como se observa en la tabla, los coeficientes son diferentes, destacando a favor del modelo de McCabe et al. el menor número de inconsistencias lógicas (3 frente a 5). Por

el contrario, la mayoría de los tests de validez predictiva indican que es ligeramente peor en este sentido el modelo de rangos que el modelo SG, lo cual no debe sorprender ya que las predicciones del modelo de Brazier et al. se comparan con los mismos datos que se utilizaron para su estimación, cosa que no ocurre en el caso del modelo de McCabe et al. El test de Ljung-Box de autocorrelación de errores sugiere que la relación entre el error de predicción y la utilidad del estado de salud observada es menos fuerte para el modelo de rangos que para el modelo SG. Por último, Los resultado muestran que el supuesto de independencia de alternativas irrelevantes no se confirma, pese a que los resultados no son consistentes a lo largo de todos los rangos.

Los resultados del modelo en términos de validez predictiva son algo peores a los obtenidos por Brazier et al. (2002), si bien esto puede explicarse en parte debido a que las utilidades observadas (halladas vía lotería estándar) que pretenden explicarse forman parte de los datos que sirvieron a Brazier et al. para estimar su modelo.

2.2.7. Propiedades empíricas del SF-6D en comparación a otras medidas

2.2.7.1. El “efecto suelo” del SF-6D

Es conocido que el SF-36 sufre del llamado “efecto suelo” (Bindman et al., 1990), expresión que hace referencia al hecho de que el sistema descriptivo no capta adecuadamente los estados de salud particularmente graves (aquellos a los que les corresponde un índice de utilidad más bajo). Esta deficiencia del SF-36 se ha trasladado al SF-6D, que resulta de redimensionar aquél, y son diversos los estudios que confirman la existencia de dicho “efecto suelo” en el SF-6D (Brazier et al., 1998, 2002, 2004b; O’Brien et al., 2003; Longworth y Bryan, 2003; Bryan y Longworth, 2005; Bharmal y Thomas, 2006; Lamers et al., 2006; Buxton et al., 2007).

Brazier et al. (2002) obtienen en su estudio de valoración de estados SF-6D con la lotería estándar una muy baja proporción de valores por debajo de cero y muy pocos por debajo de -1 (antes de reescalar); muchos menos de los que se han hallado para el HUI y el EQ-5D.¹⁷ Los valores negativos en sus valoraciones directas no llegan al 7% del total (245/3518), mientras que casi una cuarta parte de los valores (23%) se sitúa entre

¹⁷ Ya las primeras medidas del SF-6D obtenidas por Brazier et al. (1988) apuntaban que los estados definidos mediante este sistema descriptivo raramente se consideraban peores que la muerte, pese a la posibilidad que de hacer tal cosa brindaban los métodos de elicitación. Este resultado parece sugerir que el SF-6D, como el SF-36, no cubre los estados de salud “extremos” que con otros métodos de valoración se consideran peores que la muerte.

0.9 y 1. Incluso el “peor” estado de salud SF-6D fue considerado mejor que la muerte por una amplia mayoría de los entrevistados. Este dato, la proporción de sujetos que considera que el peor estado descrito por el sistema es mejor que la muerte, constituye una clara evidencia del “efecto suelo”. Dicha proporción es del 73% para el SF-6D (Brazier et al., 2002), mientras que para el HUI3 es del 9% (Feeny et al., 2002) y para el EQ-5D es del 8% (Dolan, 1997).

También se refleja el “efecto suelo” en el rango de valores o puntuaciones generado por cada instrumento de medida: el valor más bajo para el SF-6D es 0.35, en tanto que para el HUI3 es -0.36 y para la tarifa MVH-TTO del EQ-5D es -0.59. La conclusión es que, para estados graves, el EQ-5D y el HUI3 puede que sean medidas más apropiadas para generar valores que puedan utilizarse en la evaluación económica.

Lo contrario cabría decir para el caso de estados de salud leves, pues el EQ-5D parece adolecer de un “efecto techo”, ya que mucha gente describe su salud como “perfecta” con este instrumento. Esto no ocurre con el SF-6D, como se ha puesto de manifiesto en diversos estudios, como los de Petrou y Hockley (2005) para el Reino Unido con el SF-6D(SF-36), o Bharmal y Thomas (2006) para los Estados Unidos con el SF-6D(SF-12), por citar los más recientes. Otra evidencia de que el SF-6D parece ser inmune al “efecto techo” es que apenas 20 de las más de 3500 valoraciones directas obtenidas por Brazier et al. (2002) con la lotería estándar para 249 estados de salud se tradujeron en una puntuación igual a 1.

En el mismo sentido apuntan los resultados de una comparación entre las puntuaciones del EQ-5D y el SF-6D en un estudio con pacientes transplantados hepáticos (Longworth y Bryan , 2003; Bryan y Longworth, 2003): para los estados con un valor medio entre las puntuaciones de ambos instrumentos inferior a 0.4 (estados más graves), el SF-6D arroja unos valores sistemáticamente superiores a los del EQ-5D; para estados con puntuaciones más elevadas (estados más leves), el SF-6D tiende a proporcionar valores inferiores a los del EQ-5D. También Szende et al. (2004) señalan que el SF-6D proporciona utilidades más bajas que el EQ-5D en pacientes asmáticos con un buen control de la enfermedad y puntuaciones más altas que éste en pacientes que controlan mal su dolencia, siendo las utilidades similares con ambos instrumentos en el caso de pacientes de asma con un nivel de control medio. Por último, Petrou y Hockley (2005), con datos de la Encuesta de Salud de Inglaterra de 1996, obtienen puntuaciones medias más bajas en el SF-6D que en el EQ-5D para los sujetos que declaran que su salud es

“muy buena”, “buena” o “regular”, y puntuaciones más altas en el SF-6D que en el EQ-5D para los encuestados que dicen tener un estado de salud “malo” o “muy malo”.

Las comparaciones con el HUI3 arrojan conclusiones similares. Sobre una muestra de 300 pacientes de cardiología, O'Brien et al. (2003) comprueban cómo la aplicación de los algoritmos del HUI3 (Feeny et al., 2002) y del SF-6D (Brazier et al., 2002) da lugar a una discrepancia sistemática entre las puntuaciones derivadas de uno y otro instrumento. Así, para valores HUI3 mayores que 0.75, las utilidades SF-6D correspondientes son notablemente inferiores, mientras que para utilidades HUI3 por debajo de 0.4, los valores SF-6D son claramente superiores. De igual modo, los 11 pacientes a los que le era asignada una puntuación negativa con el HUI3, obtenían puntuaciones iguales o superiores a 0.25 con el SF-6D.

Entre las razones que se aducen para explicar estas discrepancias sistemáticas, O'Brien et al. (2003) señalan en primer lugar las que se deben al instrumento descriptivo en sí mismo. Así, en el caso concreto de la comparación entre HUI3 y SF-6D, los conceptos de “salud” que subyacen a ambos instrumentos son radicalmente diferentes. El HUI3 se centra en deficiencias visuales, auditivas, cognitivas o en el habla, en tanto que el SF-6D se basa en la definición más amplia defendida por la Organización Mundial de la Salud, orientada a los problemas de funcionamiento físico, mental y social.

Una segunda razón que habitualmente se ofrece para explicar las diferencias atañe al método de elicitación empleado en los estudios de valoración que sirvieron de base a la estimación de los algoritmos. En el EQ-5D se utilizó la “equivalencia temporal” (TTO), mientras que en el SF-6D se utilizó una versión modificada del SG, y son varios los estudios¹⁸ que sugieren que los valores que se obtienen con el SG tienden a ser mayores que los obtenidos con el TTO. En el HUI3 las puntuaciones SG se obtuvieron mediante una transformación potencial de las valoraciones directas obtenidas mediante la escala visual analógica. No existe, sin embargo, evidencia concluyente al respecto, pues también hay autores que atribuyen las diferencias, no tanto al método de elicitación, cuanto a su forma concreta de administración.

Por último, los algoritmos también difieren en los métodos estadísticos utilizados para la estimación de los coeficientes. Como señalan O'Brien et al. (2003), el algoritmo HUI3 se basa en la forma multiplicativa de la función de utilidad, que permite un tipo

¹⁸ Cfr. Torrance (1976), Read et al. (1984), Hornberger et al. (1992), Stiggelbout et al. (1994), Bleichrodt y Johannesson, (1997), Lenert et al. (1998) y Dolan (2000).

de interacción entre dimensiones que se supone igual entre todas ellas y para todos los niveles. Por el contrario, la “tarifa” del SF-6D se estimó mediante técnicas de regresión, eligiéndose un modelo aditivo lineal sobre la base de la bondad de ajuste y la parsimonia.

2.2.7.2. Propiedades psicométricas

Una de las principales propiedades psicométricas que se debe exigir a un instrumento de medida de la calidad de vida relacionada con la salud, como el SF-6D, es la validez empírica. En un reciente estudio con una muestra de gran tamaño (n=16443) extraída de la Encuesta de Salud de Inglaterra de 1996, Petrou y Hockley (2005), hallan evidencia de que el SF-6D es una medida de utilidad multiatributo alternativa al EQ-5D empíricamente válida y eficiente, capaz de discriminar entre indicadores externos del estado de salud. La validez empírica del SF-6D queda demostrada, en primer lugar, por la existencia de diferencias estadísticamente significativas en las puntuaciones SF-6D entre sujetos que describieron su salud como “muy buena”, “buena”, “regular”, “mala” o “muy mala”. En segundo lugar, el SF-6D generó puntuaciones (utilidades) que decrecían de forma monótona con el deterioro del estado de salud auto-percibida.¹⁹

Otro estudio reciente (Lamers et al., 2006) con pacientes aquejados de problemas de salud mental confirma la validez empírica del SF-6D, comprobándose que las utilidades medias y medianas obtenidas con este instrumento decrecen conforme aumenta la gravedad del problema de salud mental, medida a través de un instrumento de medida específico, el *Symptom Checklist* (SCL-90).

Una de las principales ventajas del SF-6D frente a otros sistemas de clasificación y, en particular, frente al EQ-5D es su mayor sensibilidad a la hora de detectar pequeños cambios en la salud de los pacientes. Esta ventaja proviene, en primer lugar, del hecho de que el SF-6D define un número mucho mayor de estados (18000) que el EQ-5D (243), de manera que algunos niveles de discapacidad captados por el SF-6D no se observan en el EQ-5D. En consecuencia, la capacidad del SF-6D para detectar mejoras en el estado de salud de pacientes que están cerca del máximo de la escala de utilidad es mucho mayor que la del EQ-5D. Téngase en cuenta que, si tomamos el algoritmo

¹⁹ Los autores advierten, no obstante, de que los tests de validez empírica se han aplicado únicamente a datos de “sección cruzada” o corte transversal, no siendo posible concluir nada acerca de la validez “longitudinal” del instrumento de medida.

MVH-TTO²⁰ del EQ-5D, encontramos un vacío entre el valor 0.883 y el valor 1, pues no existe ningún estado de salud con una utilidad comprendida en dicho intervalo. Además, la redacción de los ítems del SF-6D, que incluye aspectos positivos y negativos de la salud, puede contribuir a un mayor grado de sensibilidad a indicadores de salud externos, como también puede ayudar a ello el horizonte temporal (“las últimas 4 semanas” en el SF-6D, “hoy” en el EQ-5D).

Esta mayor sensibilidad ha sido puesta de manifiesto en diversas investigaciones (Longworth y Bryan, 2003; Petrou y Hockley, 2005). En el estudio de Petrou y Hockley (2005) el SF-6D resultó ser entre un 31% y un 100% más eficiente que el EQ-5D en detectar diferencias en el estado de salud auto-percibido, y entre un 10% y un 46% más eficiente en detectar diferencias en enfermedad, discapacidad o uso de medicamentos, lo cual se explica fundamentalmente por la mayor sensibilidad del sistema descriptivo SF-6D. Una de las implicaciones que se derivan de la mayor eficiencia del SF-6D a la hora de detectar indicadores externos del estado de salud, es el menor requerimiento en términos de tamaño muestral para detectar diferencias mínimamente importantes (MID) en los estados de salud en el curso de ensayos clínicos (Petrou y Hockley, 2005).

No obstante, Bharmal y Thomas (2006), con una muestra de 11248 individuos, concluyen que el SF-6D –concretamente el SF-6D(SF-12)– no es capaz de discriminar entre individuos que declaran disfrutar de salud perfecta según el EQ-5D pero que presentan diferentes grados de morbilidad, algo que sí consiguen hacer la puntuación resumen del componente físico del SF-12 (el PCS-12) y la escala visual del EuroQol (EQ-VAS).

Son diversas las propiedades psicométricas, además de las mencionadas, que deberían considerarse al elegir una medida concreta de CVRS basada en preferencias con fines evaluativos. En particular, convendría tener en cuenta el desempeño de los instrumentos de medida en términos de factibilidad o viabilidad y fiabilidad, así como otras formas de validez distintas de la validez empírica, como la validez teórica, la validez de constructo, etc. Es éste un ámbito en el que resulta necesario disponer de más evidencia, pues la existente es ciertamente escasa. Si acaso cabría señalar la desventaja comparativa que el SF-6D puede tener frente al EQ-5D desde el punto de la viabilidad. Así, por ejemplo, en la Encuesta de Salud de Inglaterra de 1996, el porcentaje de sujetos

²⁰ Tarifa del EQ-5D estimada a partir de valores obtenidos mediante la equivalencia temporal (TTO) por Dolan (1997).

que no completaron el cuestionario SF-36 fue más del cuádruple que el de los que no completaron el EQ-5D (10.38% vs. 2.41%). La mayor sensibilidad y eficiencia del SF-6D se vería, así, compensada por peores tasas de respuesta y de completitud.

3. Motivación del estudio.

3.1. Antecedentes

La revisión bibliográfica cuyos resultados hemos expuesto y comentado en el epígrafe anterior muestra claramente que, si bien el SF-36 es la medida genérica de CVRS más ampliamente utilizada (Garratt et al., 2002), también resulta evidente que ofrece varias limitaciones para su aplicación en el establecimiento de prioridades sanitarias. En concreto, ni las puntuaciones por dimensión ni tampoco las puntuaciones resumen del SF-36 pueden emplearse como pesos o utilidades (índices cardinales de preferencias) para el cálculo de los Años de Vida Ajustados por la Calidad (AVAC). El mayor impedimento que plantea el SF-36 radica en que, en definitiva, no es capaz de captar las preferencias de los encuestados. No al menos de la forma que el cómputo de los AVAC requiere. Tal y como Brazier et al. (1999) señalan, las reglas de agregación de las puntuaciones registradas en instrumentos del estilo del SF-36 son arbitrarias, ya que se asientan sobre la hipótesis de que la distancia en términos de preferencias entre las opciones de respuesta de un ítem determinado es constante. Esto hace que, a lo sumo, las preferencias que se declaran en estos cuestionarios sean ordinales, pero no cardinales como, insistimos, necesitamos para aplicar los AVAC en ejercicios de evaluación económica de tecnologías sanitarias.

Hemos visto cómo varios investigadores han intentado superar las limitaciones del SF-36, reteniendo al tiempo parte de su amplia riqueza descriptiva. En concreto, nuestra revisión ha permitido identificar más de una decena de algoritmos o modelos predictivos que permiten asignar utilidades a los 18000 estados de salud que puede describir el SF-36 y a los 7500 que caracteriza el SF-12. Estos algoritmos son clasificados por Pickard et al. (2005) como de “primera” o de “segunda” generación, según el algoritmo se base en medidas directas, o en medidas indirectas de las preferencias (los algoritmos estimados para otros instrumentos). En el primer caso (algoritmos de primera generación) las respuestas al SF-36/SF-12 se “transforman” en un determinado estado de salud, cuya utilidad se mide directamente por medio de técnicas como la Escala Visual Analógica o de la Lotería Estándar. En el caso de los algoritmos de segunda generación, las respuestas al SF-36/SF-12 se “traducen” a sistemas multiatributo de clasificación de estados de salud como el EQ-5D o el HUI, los

cuales generan indirectamente medidas de preferencias. De todos los algoritmos revisados, Pickard et al concluyen que aquéllos desarrollados por Brazier et al (2002, 2004) son los que poseen una base teórica y metodológica más sólida, recomendando que sean éstos los que se utilicen cuando se efectúe un análisis de sensibilidad para examinar la robustez de los resultados alcanzados.

El eslabón ideado por Brazier et al para ligar al SF-36 con una tarifa social de utilidades es, como ya hemos tenido ocasión de describir en anteriores epígrafes, el instrumento conocido como SF-6D. Esta idea es la que motiva el presente estudio. En esta memoria describimos el proceso y resultados de la estimación de varios modelos para el SF-6D, a partir de la medición directa de las preferencias de una muestra de la población general de la Región de Murcia. El algoritmo resultante permite generar pesos de CVRS susceptibles de emplearse en el cálculo de AVAC. A su vez, esta estimación permitirá el cálculo de ratios incrementales coste-utilidad (las populares ICER) que podrían informar la toma de decisiones respecto de la adopción de nuevas tecnologías.

3.2. Objetivos

El objetivo principal del proyecto es la estimación empírica de un algoritmo que asigne a cada condición o problema de salud construido a partir de las respuestas al cuestionario SF-36, un valor representativo de las preferencias de la población general murciana. Dicho algoritmo predecirá las utilidades de los estados SF-6D, de modo que tales utilidades puedan emplearse en diferentes contextos como, por ejemplo, la evaluación económica de intervenciones sanitarias, el análisis coste-beneficio generalizado del sistema murciano de salud, o simplemente el cómputo del stock de salud de la población murciana.

Los objetivos secundarios del proyecto son dos:

1. La recopilación de información cualitativa sobre la salud de la población general murciana, a partir de las puntuaciones otorgadas en el SF-36 a cada una de sus ocho dimensiones. Dicho acopio de información resultará útil para profundizar en el conocimiento de la calidad de vida de los habitantes de la región.
2. El perfeccionamiento metodológico del procedimiento de estimación del algoritmo. Las potenciales mejoras respecto a los procedimientos usuales pueden concretarse en dos nuevos elementos. En primer lugar, la utilización de una técnica de obtención de preferencias *a priori* libre de algunos de los sesgos

normalmente atribuidos al método empleado en los estudios de Brazier et al. Hablamos de una técnica que algunos autores han denominado “doble lotería” (Pinto y Abellán, 2005) o, con mayor precisión, *probability lottery equivalence* (Bleichrodt et al., 2007). En segundo lugar, junto a los modelos econométricos habituales (modelos de efectos aleatorios), se estima un modelo de carácter semiparamétrico, especificado sobre supuestos más laxos que los asumidos en la inferencia paramétrica (p.ej. normalidad) y presuntamente libre de sesgos de selección.

4. Material y métodos.

4.1. Selección de estados y métodos de valoración

Antes de exponer el modo en que se seleccionaron los estados que fueron valorados directamente por los individuos integrantes de la muestra, conviene poner de manifiesto que en esta tarea se ha de afrontar un *trade-off* entre dos objetivos: de un lado la conveniencia de maximizar el número de observaciones por estado y, de otro, el deseo de ampliar lo máximo posible el subconjunto de estados que, de los 18000 posibles, serán valorados directamente y que, en consecuencia, servirán de base para la posterior modelización de la tarifa mediante métodos de inferencia estadística.

Esta disyuntiva se plantea más allá de cual sea el tamaño muestral debido a que los métodos (“lotería estándar” en el estudio de referencia, “doble lotería” en nuestro caso) y los procedimientos (“ping-pong” en Brazier et al., 2002; PEST²¹ en nuestro estudio) tienen un elevado coste en tiempo y limitan notablemente el número de estados de salud que un mismo individuo puede valorar. En el estudio de Brazier et al. (2002) cada sujeto valoró seis estados de salud; en el nuestro el número de estados por sujeto se reduce a cinco.

A la vista del diseño del estudio de Brazier et al. (2002) cabría afirmar que éstos atribuyeron una mayor importancia al segundo de los objetivos citados más arriba, por cuanto que el total de estados de salud SF-6D objeto de valoración mediante la “lotería estándar” (SG) ascendió a 249, mientras que el número medio de observaciones por estado fue de 15, siendo algunos estados valorados por menos de 10 individuos.

En nuestro estudio hemos optado por incrementar sustancialmente el número de observaciones por estado a costa, eso sí, de reducir el tamaño del subconjunto objeto de valoración directa. Para adoptar esta decisión se ha tenido en cuenta el hecho de que la selección de 49 estados que resulta de aplicar un procedimiento de ortogonalización garantiza la estimación de una forma funcional aditiva, a lo que se añade la evidencia de que los modelos que mejor se comportan recogen básicamente los efectos directos

²¹ Ver Apéndice 1.

causados por los niveles de cada una de las dimensiones (modelos de efectos principales).²²

No obstante, puesto que es nuestra intención contrastar algunos modelos que incorporen interacciones entre dimensiones, hemos añadido a los 49 estados resultantes del procedimiento de ortogonalización (en realidad, 48 si descontamos el estado “salud normal” (*all best*), 111111, que no se valora por tener asignada, por definición, una utilidad igual a la unidad) un total de 30 estados que, a juicio del equipo investigador, recogen suficientemente la variabilidad de dimensiones y niveles del instrumento SF-6D. Entre estos 30 estados se han incluido los llamados estados “esquina” (*corner states*), definidos como aquéllos en los que solo un atributo se sitúa en su peor nivel mientras que los cinco restantes se mantienen en su mejor nivel (nivel 1), así como el peor estado (*all worst* o *‘pits’ state*), en el que todos los atributos están en su máximo (peor) nivel: 645655. Los primeros se han incluido con la finalidad de poder afrontar en una futura investigación un enfoque “descompuesto” (basado en la teoría de la utilidad multiatributo) en la estimación de la tarifa del SF-6D.²³

Tabla 14. Estados de salud valorados directamente

111131	112451	532113	242541	342623	623443
111411	121622	612321	331551	543152	245354
311112	133322	121525	514224	543233	444245
411111	141314	224152	632115	243543	524345
111115	222332	231424	641232	333633	532454
113131	235121	122255	223534	335244	644342
115111	333221	135242	333433	434631	325554
211213	641111	325412	343333	445125	434545
222222	132144	512522	423433	531435	444544
422211	132612	525311	423514	634512	445354
124123	144411	115533	431353	643233	615654
411142	322134	213615	521641	224635	545654
621121	412422	234243	314345	344425	645655

En consecuencia, se miden directamente las utilidades asociadas a 78 estados de salud (los que se muestran en la Tabla 14), valores a partir de los cuales se estimará la tarifa. Aunque el número de estados es relativamente bajo (algo menos de un tercio de los que integran la base de datos de Brazier et al.), el diseño experimental nos ha permitido, como se ha dicho, disponer de un número medio de observaciones por estado cuatro veces superior al del citado estudio (más de 60 frente a 15).

²² La razón por la que Brazier et al. (2002) añaden 200 estados a los que resultan de la ortogonalización es la intención de recoger con ellos más adecuadamente las posibles interacciones entre dimensiones.

²³ Este enfoque alternativo complementarí la estimación “compuesta” o por inferencia estadística que abordamos aquí, si bien trasciende los objetivos de la presente investigación.

Para la valoración directa del subconjunto de 78 estados de salud SF-6D se ha recurrido al método conocido como “doble lotería” o “lotería equivalente” (McCord y De Neufville, 1986). Esta técnica minimiza los sesgos generalmente reconocidos al método “encadenado” de la lotería estándar, aplicado por Brazier et al. (2002), habiendo sido contrastada su superioridad en el ámbito de la salud. Tal y como se ha advertido en varias investigaciones (p.ej. Llewellyn-Thomas et al., 1982) las utilidades encadenadas pueden resultar mayores que los valores obtenidos de modo directo, lo cual contribuiría a explicar el porqué del “efecto suelo” del SF-6D, tratado en una sección anterior de este informe. Una posible explicación para esta discrepancia entre utilidades directas y encadenadas la brindan dos teorías de utilidad no esperada conocidas como *Rank-Dependent Utility* y *Prospect Theory*. En esencia tales teorías predicen que los encuestados sobrevaloran las probabilidades pequeñas, mientras que infravaloran las probabilidades grandes, lo cual conduce a que el encadenamiento pueda generar utilidades de mayor magnitud (véase a este respecto Bleichrodt, 1996). La posible aplicación de RDU a la transformación de las utilidades de estados de salud ha sido discutida en Abellán y Pinto (2000b).

Dentro de las técnicas de “doble lotería”, el método elegido en nuestro estudio es el denominado *probability lottery equivalence* (PLE), según la terminología utilizada en Bleichrodt et al. (2007). La diferencia entre esta técnica y la denominada *value lottery equivalence* (VLE) radica en el parámetro con el que se logra la indiferencia: en el caso de la PLE dicho parámetro es la probabilidad, mientras que en la VLE es otro distinto como, por ejemplo, la duración asociada a un estado de salud.

En esencia, el método de elicitación que aplicamos pasa por encontrar la distribución de probabilidades que hace que el sujeto sea indiferente entre dos “loterías”: una en la que los posibles sucesos son la salud normal y la muerte; y otra en la que los *outcomes* son la salud normal y el estado crónico Q cuya utilidad se pretende obtener. El diseño concreto que hemos aplicado consiste en mantener constante las probabilidades de la lotería en la que el estado Q es uno de los resultados posibles y buscar la indiferencia con la probabilidad asociada a la salud normal en la lotería donde está presente la muerte como probable *outcome*. Se trata, en consecuencia, de hallar el valor de p que consigue la indiferencia entre las dos loterías²⁴:

²⁴ Ver Apéndice 1 para una descripción más detallada de los escenarios de valoración y del tipo de preguntas que se plantean a los entrevistados.

$$(SN, p, Muerte) \sim (SN, 0.5, Estado Q) \quad (9)$$

En la mayor parte de los casos se obtiene que $p > 0.5$ lo que indica que el sujeto admite un cierto riesgo de morir $(1-p) < 0.5$ a cambio de evitar un riesgo del 50% de sufrir el estado Q de manera crónica. Sin embargo, determinados estados de salud pueden llegar a ser considerados, por su especial gravedad, “peores que la muerte”, en cuyo caso se obtienen valores de p inferiores al 50%. Dado que se ha utilizado un cuestionario asistido por ordenador, y puesto que la aplicación informática incluye un procedimiento secuencial de convergencia al valor de indiferencia (ver Apéndice 1), ha resultado necesario incluir una pregunta inicial que nos permita saber si el encuestado considera el estado Q mejor o peor que la muerte (aunque el *framing* del método es idéntico en ambos casos, lo que varía en uno u otro supuesto es el sentido ascendente o descendente en el que debe variarse el estímulo p a la vista de las respuestas del sujeto).

4.2. Recogida de datos

4.2.1. La muestra

Dada la imposibilidad material, ya mencionada, de que un mismo sujeto valore todos los estados seleccionados, para la obtención de las utilidades de los 78 estados de salud se ha aplicado un diseño “entre-muestras”, esto es, los estados se han repartido en subconjuntos de menor tamaño que se distribuyen para su valoración por subgrupos de la muestra total. De manera análoga a lo realizado en el estudio de referencia, donde cada sujeto valoró un total de seis de los 249 estados seleccionados, el subconjunto de estados que cada sujeto tendrá que valorar en nuestro estudio será de cinco.

A la hora de “repartir” los estados entre los sujetos, una posibilidad podría haber sido la de aplicar un procedimiento aleatorio, si bien hemos considerado preferible definir subconjuntos de 5 estados y asignar cada subconjunto a una parte concreta de la muestra. Con este proceder se pretende, en primer lugar, garantizar que todos los sujetos valoran estados de diferente gravedad (leves, moderados, graves, ...). A tal fin, los estados se ordenaron atendiendo a la suma no ponderada de los niveles de sus seis dimensiones, agrupándolos en cinco categorías atendiendo a esta pseudo-ordenación, y asignando a cada subconjunto un estado procedente de cada uno de estos cinco grupos.²⁵

²⁵ Determinados estados de salud (los “estados esquina” y el peor estado que describe el SF-6D) fueron asignados a dos subconjuntos diferentes siendo, en consecuencia, valorados por el doble de sujetos que el resto.

En segundo lugar, al adscribir cada subconjunto de 5 estados a una submuestra pretendemos mantener la representatividad por sexo y cuotas de edad en todos y cada uno de los estados valorados, algo que no sería posible si la asignación de los estados (o subconjuntos de éstos) a los sujetos fuese aleatoria. Así, la muestra total está, de hecho, compuesta por 17 submuestras ($n_i=60$), todas ellas representativas de la población adulta por sexo y grupos de edad.

4.2.2. El cuestionario

El cuestionario se estructuró en cuatro partes, la primera de las cuales se dedicó a explicar a los sujetos las características del sistema de clasificación de estados de salud SF-6D a través de un “tutorial” en el que, a partir de un ejemplo, se exponían las diferentes dimensiones del instrumento y el significado de los niveles asociados a cada dimensión. Describimos a continuación, brevemente, el contenido de las tres secciones restantes del cuestionario, que han servido para obtener los resultados que aquí se presentan. Una exposición más detallada puede encontrarse en el Apéndice 1.

En la parte 2, una vez que el encuestado está familiarizado con el instrumento de clasificación, se le mostró la descripción en términos del SF-6D de los cinco estados de salud que había de valorar, etiquetados como *V*, *W*, *X*, *Y*, *Z*. Para facilitar la comprensión por parte de los entrevistados, se utilizaron ayudas visuales en forma de “tarjetas”, en las que la descripción asociada a cada dimensión se acompañaba del número identificativo del nivel en cuestión, resaltado sobre un gradiente de color que iba del verde (menos grave) al rojo (más grave). La primera tarea que debía afrontar el entrevistado era la de situar cada uno de los cinco estados sobre un “termómetro” rotulado de 0 (*el peor estado de salud imaginable*) a 100 (*el mejor estado de salud imaginable*). Sobre el mismo termómetro debía el encuestado marcar la alternativa *muerte inmediata*, que podía coincidir o no con el extremo inferior del intervalo, así como situarse por encima de alguno de los estados sometidos a valoración, en cuyo caso se habría de interpretar que el sujeto consideraba dicho estado “peor que la muerte”. Tras situar los estados sobre la escala, el sujeto tenía la posibilidad de revisar sus valoraciones, modificando el valor concreto de cada estado –o incluso la propia ordenación de los mismos–.

La parte tercera de la entrevista es la que incluye el método elegido para la obtención de utilidades, un escenario de “doble lotería” que denominamos *probability lottery*

equivalence (PLE), que se repetía consecutivamente cinco veces, una por cada uno de los estados de salud que le correspondía valorar al sujeto, siendo aleatorio el orden de aparición de los estados V al Z dentro de esta sección.

En esencia, el procedimiento PLE que aplicamos consistió, como se ha dicho en el apartado 4.1 de este informe, en encontrar el valor de p que consigue la indiferencia entre las dos loterías:

$$(SN, p, Muerte) \sim (SN, 0.5, Estado Q) \quad (10)$$

La pregunta inicial fijaba $p=0.5$, con el fin de conocer si el sujeto consideraba el estado de salud mejor o peor que la muerte. Si el sujeto elegía la lotería de la derecha, con ello manifestaba que, a igualdad de probabilidades de disfrutar de una salud normal (50% en ambos casos), prefería la alternativa en la que el riesgo se concretaba en sufrir un deterioro en la salud con carácter crónico (el *estado Q*) antes que afrontar el riesgo de morir. A partir de aquí, sucesivas elecciones con distintos valores de $p > 0.5$ conducían a encontrar el valor de indiferencia.

Si para una probabilidad de morir del 50% elegía la primera lotería, indicaba con su respuesta que el estado en cuestión era tan indeseable que prefería arriesgarse a morir que a sufrir dicho estado de manera crónica. En tal caso, los valores de p que se presentaban en sucesivas elecciones eran siempre inferiores a 0.5, tratando de buscar con ello el mínimo valor de la probabilidad de “salud normal” que se estaba dispuesto a afrontar con tal de evitar un riesgo del 50% de sufrir el *estado Q*.

La cuarta parte del cuestionario se destinó a recoger información sobre el estado de salud del propio encuestado –no ya sobre sus preferencias respecto a hipotéticas situaciones de salud, que era el objeto de las partes 2 y 3–, así como acerca de sus características personales. Para acceder a la descripción y valoración del propio estado de salud se utilizaron tres instrumentos: el “auto-clasificador” EQ-5D, el cuestionario de salud SF-36 (versión 2) y una escala visual en forma de termómetro, sobre la que el sujeto debía valorar su estado de salud en relación con los extremos (*el mejor estado imaginable–el peor estado imaginable*), pero también respecto de los cinco estados y la alternativa *muerte inmediata* que el mismo encuestado había evaluado en la parte 2 del cuestionario.

4.2.3. Cálculo de las utilidades

Si identificamos por p^* al valor que consigue la indiferencia en la expresión (10), siendo $U(SN)=1$ y $U(Muerte)=0$, y asumiendo que se verifican los supuestos de la utilidad esperada, podemos deducir de la citada expresión que $p^* + 0 = 0.5 + 0.5 U(Q)$, de donde se obtiene el siguiente valor para la utilidad del estado de salud Q :

$$U(Q) = 2p^* - 1 \quad (11)$$

El cálculo de la utilidad del estado de salud valorado con el método PLE es idéntica con independencia de que el estado se considere mejor o peor que la muerte. Una ventaja adicional de la forma elegida para administrar el método de elicitación (al margen de la ya argumentada superioridad del procedimiento frente a la lotería estándar), radica en que las utilidades que se obtienen para estados peores que la muerte están acotadas entre 0 y -1 , no siendo necesario aplicar método alguno de reescalamiento, con lo que se evita la arbitrariedad que este tipo de procedimientos conlleva en mayor o menor grado.

4.2.4. Cálculo de las puntuaciones resumen del SF-36

El cuestionario de salud SF-36 está compuesto por 36 ítems que pretenden recoger todos los aspectos relevantes para caracterizar la salud de un individuo. Con estas preguntas se trata de cubrir, al menos, 8 aspectos o dimensiones: Función Física; Rol Físico; Dolor Corporal; Salud General; Vitalidad; Función Social; Rol Emocional y Salud Mental. Para cada una de estas dimensiones se pueden computar escalas de puntuación, fácilmente interpretables, caracterizadas todas ellas por encontrarse ordenadas, de tal suerte que cuanto mayor es el valor obtenido mejor es el estado de salud. En la Tabla 15 se ilustra el significado de los valores extremos para las distintas escalas.

El cómputo de las escalas arriba mencionadas exige la realización de tres pasos consecutivos. En primer lugar, resulta necesario un proceso de homogeneización, consistente en recodificar las respuestas otorgadas por los entrevistados, de manera que los valores numéricos sean congruentes con la regla antes citada de que una mayor puntuación equivale a disfrutar de un mejor estado de salud. Esto es preciso porque en algunas dimensiones (p.ej. Dolor Corporal) la ordenación inicialmente obtenida va, exactamente, en sentido contrario. En segundo lugar, se han de computar las escalas de puntuación “crudas” (“*raw-scores*”) para cada una de las dimensiones. Este procedimiento no requiere más que agregar los diversos valores numéricos previamente

recodificados. En tercer lugar se realiza una transformación lineal de estas escalas con la finalidad de reescalar las puntuaciones en el intervalo abierto 0-100.²⁶

Tabla 15. Dimensiones, número de ítems y valores extremos del SF-36

Dimensión	Nº ítems	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.
Rol Físico	4	Tiene problemas con el trabajo u otras actividades diarias a causa de su salud física.	No tiene problemas con el trabajo u otras actividades diarias a causa de su salud física.
Dolor Corporal	2	Sufre dolor muy intenso y extremadamente limitante.	No sufre dolor.
Salud General	5	Evalúa como mala la propia salud y cree posible que empeore.	Evalúa su salud como excelente.
Vitalidad	4	Se siente cansado y exhausto todo el tiempo.	Se siente lleno de energía todo el tiempo.
Función Social	2	Problemas físicos o emocionales interfieren de forma extrema en las actividades sociales normales.	Lleva a cabo actividades sociales normales sin ninguna interferencia por motivos de salud física o emocional
Rol Emocional	3	Problemas emocionales generan problemas en el trabajo u otras actividades.	No hay problemas emocionales que impidan el normal desempeño en el trabajo u otras actividades.
Salud Mental	5	Se siente angustiado y deprimido todo el tiempo.	Se siente feliz, tranquilo y calmado, todo el tiempo.
Ítem de transición	1	Cree que su salud es mucho peor ahora que hace un año.	Cree que su salud es mucho mejor ahora que hace un año.

Fuente: Adaptado de Vilagut et al (2005).

Además de las puntuaciones en escala 0-100, resulta común el cálculo de otras escalas basadas en normas, cuya principal ventaja es que los resultados son directamente comparables con una población de referencia. Un valor normalizado superior a 50 indica una mejor condición de salud que la media de la población de referencia; lo contrario si es inferior a 50. De lo que se trata, obviamente, es de tener unos valores

²⁶ Podemos ilustrar este procedimiento de cómputo con un ejemplo aplicado a la dimensión “Dolor Corporal”, que incluye dos ítems. Imaginemos que ante la pregunta “¿Tuvo dolor en alguna parte del cuerpo durante las 4 últimas semanas?”, un entrevistado responde: “Sí, muy poco” (respuesta codificada con un 2 sobre 6 niveles). Supongamos en el segundo ítem de esta dimensión, a la pregunta “Durante las 4 últimas semanas, ¿hasta qué punto el dolor le ha dificultado su trabajo habitual (incluido el trabajo fuera de casa y las tareas domésticas)?” el mismo sujeto contesta “Un poco” (codificada con un 2 sobre 5). El primer paso será recodificar las respuestas obtenidas para que sea cierto que cuanto mayor sea el valor numérico de la respuesta, mejor es el estado de salud del encuestado. En el ejemplo ello supondría convertir el 2 (sobre 6) de la primera pregunta en un 5, y el 2 (sobre 5) de la segunda pregunta en un 4. El segundo de los pasos, consiste en la agregación de las respuestas obtenidas. En este caso, la puntuación “cruda” tomaría un valor de 9 (5+4). Para la obtención de la puntuación final en la dimensión “Dolor Corporal” sólo media realizar la transformación lineal que asegure que nuestro valor queda comprendido en el intervalo 0-100. En el ejemplo, tendríamos que restar 2, dividir por 10, y multiplicar por 100, con lo que el valor finalmente obtenido sería de 70.

poblacionales de referencia con los que comparar. En el caso español, aún no se dispone de valores generalmente aceptados, por eso recurriremos a los valores de la población norteamericana, que son los valores poblacionales de referencia más comúnmente utilizados.²⁷

Como se ha señalado en el epígrafe 2.2.2, el proyecto *International Quality of Life Assesment (IQOLA)*²⁸ demostró, además, que desde las 8 escalas anteriores se pueden extraer dos componentes principales de salud: la componente resumen física (PCS) y la componente resumen mental (MCS). Para el cálculo de estas dos medidas resumen, es necesario, de nuevo, seguir un procedimiento ordenado. En primer lugar, estandarizar (con las medias y desviaciones típicas poblacionales) las 8 escalas del SF-36. En segundo lugar, reponderar las escalas aplicando los pesos factoriales que se obtuvieron en el estudio de componentes principales. En tercer lugar, agregar y transformar para que la distribución tenga una media de 50 y una desviación típica de 10 en la población general. En la siguiente tabla quedan recogidos los valores medios y las desviaciones típicas de la población norteamericana utilizada, así como los pesos de ponderación, para cada una de las dimensiones.

Tabla 16. Parámetros utilizados para el cálculo de la componente resumen física (PCS) y la componente resumen mental (MCS) del SF-36

	Media	Desviación Típica	Pesos PCS	Pesos MCS
Función Física	83.29084	23.75883	0.42402	-0.22999
Rol Físico	82.50964	25.52028	0.35119	-0.12329
Dolor Corporal	71.32527	23.66224	0.31754	-0.09731
Salud General	70.84570	20.97821	0.24954	-0.01571
Vitalidad	58.31411	20.01923	0.02877	0.23534
Función Social	84.30250	22.91921	-0.00753	0.26876
Rol Emocional	87.39733	21.43778	-0.19206	0.43407
Salud Mental	74.98685	17.75604	-0.22069	0.48581

En el Apéndice 2, quedan recogidos los valores obtenidos en este estudio, tanto en lo que se refiere a las escalas 0-100 para las 8 dimensiones, como las relativas a las medidas normalizadas y los componentes resumen.

²⁷ En estudios en otros países se ha obtenido que las correlaciones entre los valores de referencia de los mismos y los valores de referencia norteamericanos son realmente elevadas (0.98), de tal forma, que el empleo de los valores norteamericanos en lugar de los del propio país, no debería introducir desviaciones significativas.

²⁸ Bullinger et al (1998).

4.3. Métodos de estimación de la tarifa SF-6D

4.3.1. Modelos de efectos aleatorios

Los procedimientos econométricos utilizados en la literatura para estimar tarifas a partir de sistemas descriptivos multiatributo como el SF-6D son eminentemente paramétricos. En concreto, el modelo de efectos aleatorios es el que predomina en este tipo de aplicaciones y es, por tanto, el punto de partida de nuestro estudio.

La idoneidad del modelo de efectos aleatorios se justifica por el diseño experimental predominante en este tipo de estudios y por la especificación de la ecuación a estimar en los mismos. Respecto al diseño experimental, en la mayoría de los estudios, incluido el que nos ocupa, cada encuestado valora una serie de estados de salud y, por tanto, la eficiencia de las estimaciones puede incrementarse, respecto a una estimación por Mínimos Cuadrados Ordinarios, utilizando estimadores que reconocen al encuestado como una posible fuente de variación de la variable dependiente (efectos individuales).

En segundo lugar, la práctica habitual en la literatura es explicar la utilidad asignada a un estado de salud a partir de una serie de variables indicadoras de los niveles de gravedad de las distintas dimensiones que lo describen. Esto es, los regresores son exógenos al propio encuestado por cuanto son descriptores de un hipotético estado de salud y, por tanto, ortogonales al potencial efecto individual inobservable. Por esta razón, el modelo de efectos aleatorios es preferible en términos de eficiencia a otro modelo que también permite corregir efectos individuales: el modelo de efectos fijos.

El primer modelo de efectos aleatorios que consideramos es el denominado “modelo de efectos principales”, consistente en regresar la variable dependiente (utilidad) sobre un conjunto de variables indicadoras de la presencia de niveles superiores al de referencia (unidad) en las diferentes dimensiones consideradas. Los coeficientes de interés se identifican estimando por mínimos cuadrados generalizados o por máxima verosimilitud la siguiente ecuación:

$$h_{DLi} = \sum_{d \in D} \sum_{l \in L} \beta_{dl} x_{dl} + \eta_i + \varepsilon_{DLi} \quad (12)$$

donde d y l indican la dimensión y el nivel de severidad de cada dimensión, D y L son el conjunto de valores posibles de las indicadoras d y l , e i hace referencia al encuestado. El término de error se descompone en un componente individual inobservable (η_i) que

recoge los efectos individuales y en una perturbación específica de cada estado de salud e individuo que suponemos normalmente distribuida.

La constante representa la utilidad en salud perfecta (nivel uno en todas las dimensiones), que suponemos igual a la unidad. Este supuesto, inocuo a la estimación del modelo al introducir un simple efecto de escala, permitir comparar, al mantener la misma escala, los resultados obtenidos en este trabajo con los existentes en la mayoría de los estudios previos en los que se restringe el valor de la constante a la unidad.²⁹

Asimismo, estimamos versiones aumentadas de la ecuación anterior que incluyen variables indicadoras de la presencia del nivel más grave en al menos una de las dimensiones e interacciones entre las variables indicadoras del modelo de efectos principales. El objetivo es encontrar la especificación que maximiza la bondad de ajuste con los datos obtenidos al tiempo que permite derivar tarifas consistentes (la utilidad se reduce con la gravedad) y maximiza los grados de libertad disponibles en la estimación, esto es, incorporamos un criterio de parsimonia en la elección de la especificación óptima.

4.3. 2. Enfoque semiparamétrico

Las estimaciones obtenidas a partir de modelos de efectos aleatorios adolecen de dos sesgos potenciales. En primer lugar, el derivado de los supuestos distribucionales inherentes al modelo. En segundo término, la posible presencia de sesgos de selección en las submuestras de encuestados comparadas para estimar los distintos coeficientes del modelo.

Los sesgos de selección surgen cuando determinadas características individuales que afectan de forma sistemática a la valoración que realizan los encuestados de los estados de salud se encuentran desigualmente distribuidas entre los colectivos que valoran distintos niveles de una misma dimensión y/o distintas dimensiones.

Algunos trabajos previos han señalado que determinadas características sociodemográficas como la edad, el sexo, el estado civil o el ingreso disponible afectan a la utilidad que los encuestados asignan a un mismo estado de salud (Dolan y Roberts,

²⁹ Hay razones teóricas para mantener esta restricción. Brazier et al. (2002) estiman un modelo similar a partir de respuestas individuales a una lotería estándar, cuyo valor para cada estado se derivó de acuerdo con la teoría de la utilidad esperada, donde salud plena deriva utilidad 1 y la muerte la utilidad cero. Un cambio en el valor de la constante en ese modelo implica un cambio de toda la escala y, por tanto, la no comparabilidad directa con otros modelos estimados donde la constante tome otro valor.

2002; Kharroubi et al., 2007a). Así, la convención de garantizar la representatividad de las muestras seleccionadas en términos del sexo y la edad (en intervalos) del encuestado se muestra insuficiente para garantizar a priori la irrelevancia de los sesgos de selección, especialmente en aquellos diseños experimentales, como el nuestro y la mayoría de los existentes en la literatura, en los que cada encuestado valora un subconjunto de los estados de salud seleccionados.

Una solución paramétrica a este problema pasa por incluir en un modelo de efectos aleatorios información sociodemográfica de los encuestados para, en caso de confirmarse su significatividad, controlar por los efectos composición o selección. No obstante, la situación se complica si, tal y como señala la evidencia disponible, algunas de las variables en las que no se garantiza la representatividad muestral resultan ser condicionantes relevantes de la valoración que hacen los encuestados de los estados de salud. En este último escenario sería necesario no solo controlar por dichas variables en la estimación de la tarifa, sino recuperar la distribución de las mismas a nivel poblacional para garantizar la representatividad de la tarifa estimada y, por tanto, de las conclusiones obtenidas. Esta cuestión, tan relevante a la hora de juzgar la validez de los resultados obtenidos como el diseño muestral, no ha sido abordada hasta el momento en la literatura.

Diseñar una muestra que sea representativa de la población de referencia en todas las dimensiones potencialmente relevantes al estudio sería poco menos que imposible. Nuestra propuesta es más sencilla y eficiente. En primer lugar, diseñamos la muestra garantizando la representatividad en las dimensiones que a priori se antojan más relevantes (tradicionalmente el sexo y la edad). En segundo término, recabamos información sociodemográfica potencialmente relevante y, finalmente, utilizamos los desarrollos teóricos de la literatura econométrica de “efectos de tratamiento” para, una vez comprobada la relevancia de dichas características, recuperar la distribución de las mismas en la población de referencia, eliminando los sesgos de selección. Asimismo, pretendemos realizar esta innovación con técnicas semiparamétricas que permitan obtener conclusiones válidas a nivel poblacional sin hacer supuesto alguno acerca de la distribución que sigue la utilidad que los individuos derivan de los distintos estados de salud.

El punto de partida de este procedimiento econométrico lo encontramos en Abadie (2005). La idea que subyace a esta técnica es similar a la que encontramos en

numerosos estudios bioestadísticos o epidemiológicos en los que el efecto de un determinado tratamiento se identifica comparando de forma adecuada a los grupos de tratamiento y control. Así, el efecto que un determinado nivel de gravedad en una determinada dimensión tiene en la utilidad se identifica siguiendo un procedimiento bietápico. En primer lugar, comparamos la distribución de características individuales relevantes³⁰ entre los encuestados que han valorado estados con ese nivel de gravedad en esa dimensión y aquellos que han valorado estados de salud caracterizados por la ausencia de problemas de salud (nivel uno) en esa dimensión.³¹ La comparación se realiza, para evitar los problemas de dimensionalidad derivados del número de características individuales consideradas, utilizando un simple modelo *logit*.

Los desarrollos teóricos de la literatura de efectos de tratamiento permiten afirmar que igualar la distribución de la probabilidad estimada (*propensity score*) entre “tratados” (encuestados que valoran el nivel de gravedad y dimensión cuyo efecto pretendemos determinar) y “controles” (encuestados que valoran estados de salud con nivel uno en dicha dimensión) equivale a igualar la distribución de todas las características individuales incluidas en la estimación de dicha probabilidad.

El efecto de interés puede entonces identificarse como una simple media ponderada de la valoración que realizan los encuestados de los estados de salud, donde las respuestas de los “controles” son ponderadas utilizando la probabilidad estimada en la primera etapa. Este procedimiento de ponderación garantiza que la distribución de características observables relevantes en la muestra de controles es la misma que en la muestra de tratados y, por tanto, permite realizar comparaciones libres de efectos composicionales. El único supuesto distribucional que subyace a este procedimiento es el inherente al modelo *logit* estimado en la primera etapa para comparar los dos colectivos que permiten identificar el efecto de interés.

Para clarificar la exposición denotamos por X al conjunto de características individuales relevantes a la determinación de la valoración que hace el encuestado y por Y al conjunto de valoraciones realizadas por los encuestados. Asimismo, la variable indicadora T tomará el valor 1 para los “tratados” y 0 para los “controles”. El *propensity*

³⁰ Las características individuales relevantes son aquellas que determinan de forma significativa las valoraciones de los distintos estados de salud realizadas por los encuestados. Este análisis se realiza utilizando un modelo de efectos aleatorios.

³¹ Por analogía a los coeficientes estimados con el modelo de efectos aleatorios. El coeficiente asociado a un determinado nivel de gravedad en una determinada dimensión mide, en una especificación como la inicialmente considerada, la pérdida de utilidad derivada de no estar en salud perfecta en dicha dimensión.

score puede denotarse por $p(T=1/x) \equiv p(x)$, mientras que el efecto de interés α puede escribirse como:

$$\alpha = E[Y/X, T=1] - E[Y/X, T=0] \quad (13)$$

Abadie (2005) demuestra que dicho efecto puede estimarse como:

$$\alpha = \frac{1}{P(T=1)} E \left[Y \frac{T - p(x)}{(1 - p(x))} \right] = \frac{1}{P(T=1)} E[Y\omega] \quad (14)$$

donde ω representa la función de ponderación que permite corregir la distribución de características individuales relevantes. El procedimiento de ponderación utilizado reduce el peso de aquellos “controles” con características relevantes sobrerrepresentadas respecto a la muestra de “tratados” e incrementa el peso de las respuestas dadas por aquellos controles con características relevantes infrarepresentadas respecto a la muestra de “tratados”.

Los efectos así identificados pueden no ser comparables entre sí, ya que su magnitud depende de la distribución de características relevantes en cada muestra de “tratados”. Para corregir este efecto distorsionador utilizamos una única distribución de características relevantes como referencia en todas las comparaciones realizadas. Así, podemos tomar como distribución de referencia la observada a nivel poblacional en una base de datos representativa y que incluya la información adecuada como, por ejemplo, el “Panel de Hogares de la Unión Europea”, que presenta abundante información sociodemográfica y laboral de la población española.

Los estudios desarrollados a partir de muestras en los que cada encuestado valora todos los estados de salud analizados eliminan la posibilidad de sesgos de selección entre los colectivos comparados a costa de incurrir en un importante sesgo de agotamiento del encuestado. El procedimiento de estimación implementado en el presente estudio permite evitar dicho sesgo de agotamiento ya que recupera la comparabilidad en términos de características individuales relevantes independientemente del diseño utilizado. La significatividad de los efectos estimados mediante este procedimiento se analiza utilizando los desarrollados asintóticos establecidos en Méndez (2008).

Finalmente, la inclusión de características individuales en el modelo pone en tela de juicio la superioridad (mayor eficiencia) del modelo de efectos aleatorios frente al

modelo de efectos fijos, al abrir el debate acerca de la endogeneidad de los regresores introducidos. Abordaremos esta cuestión utilizando el test de Hausman, esto es, comparando los coeficientes estimados mediante los dos procedimientos y eligiendo el estimador más eficiente en cada caso.

5. Resultados.

5.1. Participantes en el estudio

Se entrevistó a un total de 1020 individuos representativos por cuotas de edad y sexo de la población adulta, distribuidos en 17 submuestras de igual tamaño cada una (60). Las entrevistas se realizaron en el territorio de la Comunidad Autónoma de la Región de Murcia entre el 5 y el 30 de marzo de 2008. Los cuestionarios, asistidos por ordenador, fueron administrados en el domicilio de los encuestados por personal adiestrado por el equipo investigador. La duración media de las entrevistas fue de 19 minutos y 30 segundos (mediana 18 minutos y 10 segundos).

Del total de sujetos encuestados, 14 fueron excluidos del análisis posterior por incurrir en alguna inconsistencia al valorar los estados de salud, de manera que un estado “lógicamente peor” recibía una valoración superior a un estado “lógicamente mejor”. De estos 14 sujetos, 4 incurrieron en alguna inconsistencia lógica al ordenar y puntuar los estados sobre la escala visual, mientras que para los 10 restantes las inconsistencias afloraron en la respuestas al método de “doble lotería” (PLE).

Otros 15 sujetos fueron excluidos de la muestra final debido a que, en la parte del cuestionario donde se elicitaban las utilidades con el PLE, se mostraron renuentes a aceptar un mínimo riesgo de muerte en al menos tres de los cinco estados que les correspondió valorar.³² No hubo lugar a excluir ninguna encuesta incompleta, debido a que el diseño de la aplicación informática impedía el registro de la encuesta si no se cumplimentaba en su totalidad.

Por lo que respecta a la incidencia de los criterios de exclusión sobre los subgrupos en los que se dividió la muestra al objeto de “repartir” los estados de salud, si bien resultó dispar, no hay ningún estado de salud para el cual se disponga de un número de observaciones inferior a 53.

³² Los casos en los que el individuo no acepta riesgo alguno de morir con tal de evitar el estado crónico Q dan lugar a que la utilidad del estado sea igual a 1. Esto ocurre cuando el sujeto escoge siempre la lotería en la que los *outcomes* son la salud normal ($p=0.5$) y el estado crónico Q , por muy elevada que sea la probabilidad que se ofrece de recuperar la salud normal en la lotería que incluye la muerte como posible resultado. Si esto ocurre tres o más veces para un mismo sujeto, interpretamos que no es capaz de discriminar entre estados de salud según su gravedad y optamos por su exclusión.

En definitiva, la muestra que se utiliza en el análisis subsiguiente está compuesta por 991 individuos. En la Tabla 17 se presentan las características de la muestra final que, teniendo en cuenta la baja tasa de exclusión (2.8%), apenas difieren de las de la muestra original (n=1020).

Tabla 17. Características de la muestra

	<i>n=991</i>	%
<i>Sexo</i>		
Mujer	490	49.45
Hombre	501	50.55
<i>Edad</i>		
De 18 años a 29 años	217	21.90
Entre 30 años y 41 años	260	26.24
Entre 42 años y 53 años	227	22.91
Entre 54 y 65 años	170	17.15
Más de 65 años	117	11.81
<i>Estado civil</i>		
Soltero	330	33.30
Casado o con pareja de hecho	596	60.14
Separado, divorciado o viudo	65	6.56
<i>Nivel de estudios</i>		
Sin estudios	109	11.00
Primarios	235	23.71
Secundarios	340	34.31
Superiores	307	30.98
<i>Nivel de ingresos</i>		
Hasta 900 euros	75	7.57
De 901 a 1.500 euros	153	15.44
De 1.501 a 2000 euros	280	28.25
De 2.001 a 3.000 euros	297	29.97
Más de 3.000 euros	186	18.77
<i>Fumador (%)</i>	27.04	
<i>Seguro médico privado (%)</i>	20.48	
<i>Estado de salud autopercebido (EQ-5D)</i>		
11111	601	60.65
11121	157	15.84
11112	43	4.34
Otros	92	19.17
<i>Estado de salud autopercebido (SF-6D/SF-36)</i>		
111122	59	5.95
111112	42	4.24
111222	31	3.13
111111	29	2.93
Otros	859	86.68

La parte final de la Tabla 17 confirma lo señalado en otra parte de esta memoria en relación con el “efecto techo” del EQ-5D y la mayor sensibilidad del SF-6D para discriminar entre estados de salud (en particular, entre estados leves). Más del 60% de los sujetos encuestados definió su estado de salud como el mejor posible según el EQ-5D, mientras que apenas un 3% respondió al SF-36 de manera que su estado de salud quedase descrito como el 111111 del SF-6D. Este estado no es, además, el más

frecuente en la muestra, pues casi el doble de individuos describe su estado de salud como equivalente al 111122. Como puede comprobarse en la tabla, un 80% de los encuestados se agrupa en tres únicos estados de salud EQ-5D, frente a los casi 200 estados SF-6D necesarios para describir a idéntico porcentaje de la muestra.

5.2. Utilidades de los estados de salud

En la Tabla 18 se ofrecen los estadísticos descriptivos de las utilidades de una selección de 30 estados de salud.

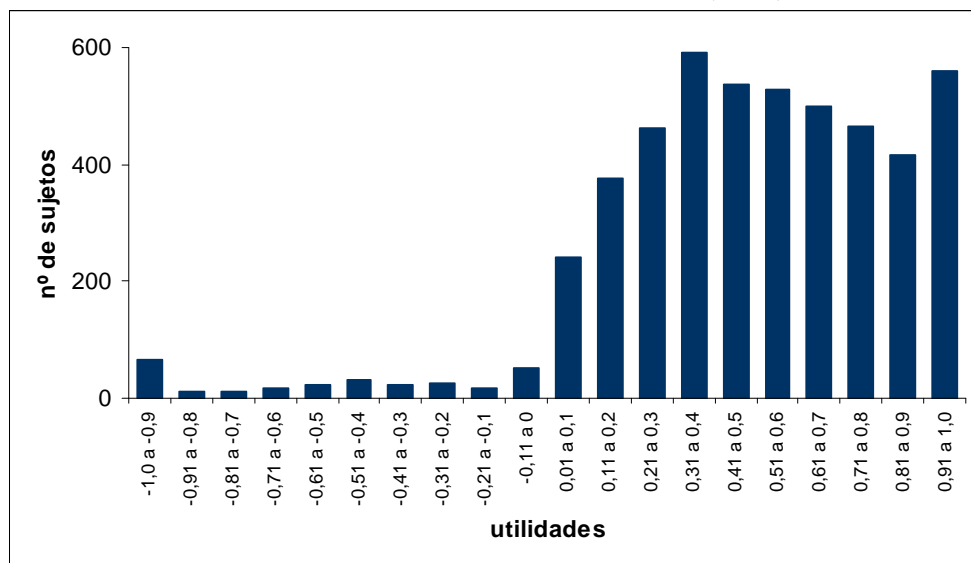
Tabla 18. Estadísticos descriptivos para 30 valoraciones de estados de salud SF-6D

Estado	n	Mínimo	Máximo	Media	Mediana	SD
113131	58	0,800	1,000	0,988	1,000	0,036
222222	60	0,520	1,000	0,891	0,930	0,130
115111	117	0,500	1,000	0,847	0,860	0,122
222332	58	0,400	1,000	0,826	0,820	0,136
111411	119	0,500	1,000	0,796	0,780	0,110
141314	56	0,240	1,000	0,755	0,800	0,176
132612	58	0,300	0,980	0,705	0,700	0,171
133322	59	0,360	1,000	0,671	0,660	0,136
621121	60	0,240	0,980	0,657	0,700	0,176
223534	54	0,240	1,000	0,641	0,600	0,181
235121	59	0,400	1,000	0,610	0,600	0,134
412422	60	0,300	1,000	0,599	0,600	0,178
121525	60	0,160	0,900	0,569	0,600	0,145
325412	58	0,000	0,900	0,552	0,510	0,182
112451	60	0,200	0,800	0,515	0,500	0,137
512522	59	0,200	0,800	0,476	0,500	0,130
122255	59	0,200	0,900	0,469	0,460	0,167
121622	60	0,180	0,700	0,451	0,460	0,103
634512	54	0,100	0,700	0,423	0,400	0,132
224152	60	0,100	0,940	0,411	0,400	0,158
314345	59	0,200	0,800	0,395	0,400	0,122
445125	60	0,100	0,800	0,369	0,330	0,140
643233	60	0,060	0,700	0,315	0,300	0,178
524345	59	-0,400	0,700	0,285	0,300	0,142
434545	59	0,060	0,520	0,241	0,220	0,102
344425	60	0,060	0,580	0,160	0,160	0,085
532454	53	-0,980	0,860	0,109	0,200	0,418
644342	58	-0,980	0,660	-0,013	0,060	0,384
615654	60	-0,960	0,620	-0,263	-0,310	0,346
645655	114	-0,980	0,500	-0,515	-0,600	0,430

Cada estado fue valorado una media de 64 sujetos, con un mínimo de 53 y un máximo de 119.³³ Los valores medios abarcan un rango que va de -0.515 a 0.988. Hasta tres estados presentan utilidades medias negativas, lo que contrasta con los resultados de Brazier et al. (2002), cuyos valores medios en todos los casos resultaron ser mayores que cero. En la mayoría de los estados (42/78) las medianas son superiores a las medias, lo que refleja una cierta asimetría en los datos agregados por estado.

A nivel individual el grado de asimetría es más evidente, como puede verse en el histograma de los valores individuales que se muestra en el Gráfico 1. El porcentaje de observaciones con signo negativo es relativamente bajo (5%), incluso inferior al 7% del estudio de Brazier et al. Sin embargo, los valores negativos obtenidos en nuestro estudio son, en términos generales, mayores en valor absoluto que los de Brazier et al. Además, una proporción significativa de los estados de salud (25/78) fueron considerados peores que la muerte (utilidad<0) por alguno de los sujetos encuestados.

Gráfico 1. Histograma y estadísticos descriptivos de las valoraciones de los estados de salud mediante la “doble lotería” (PLE)



Si comparamos la distribución de nuestras valoraciones con la de Brazier et al., podemos comprobar cómo nuestra distribución está desplazada hacia la izquierda respecto a la del estudio de referencia (ver Tabla 19). La media y la mediana son más bajas en nuestro estudio (0.508 y 0.50 frente a 0.5417 y 0.65 respectivamente), siendo mayor el grado de asimetría negativa en nuestra distribución (-1.26 frente a -0.78). Por

³³ Como se ha señalado en otra parte, los denominados “estados esquina” (*corner states*) y el peor estado del SF-6D (*pits state*) fueron valorados por dos submuestras con el fin de ser utilizados en la estimación del algoritmo mediante el enfoque descompuesto basado en la teoría de la utilidad multiatributo. Este proyecto, no obstante, queda fuera del ámbito de la presente investigación.

otra parte, nuestros valores presentan una menor dispersión, medida a través de la desviación estándar y, además, el coeficiente de curtosis indica que nuestras observaciones se concentran en mayor grado alrededor de los valores centrales que en el caso del estudio pionero.

Algunos datos en particular permiten entender las diferencias entre nuestros resultados y los de Brazier et al. Así mientras que en su estudio un 73% de los encuestados otorgó valoraciones positivas al peor estado de salud (*worst state*) descrito por el SF-6D, el 645655, un porcentaje similar (75%) de los individuos encuestados en nuestro estudio le asigna utilidades inferiores a -0.30.

Tabla 19. Estadísticos descriptivos de las valoraciones SG de Brazier et al. (2002) y de las valoraciones PLE obtenidas en nuestro estudio

	Brazier et al. (2002)	Nuestro estudio
N	3518	4955
Media	0.5417	0.5083
Mediana	0.6500	0.5000
Máximo	1.0000	1.0000
Mínimo	-0.9800	-0.9800
Desv. Est.	0.3881	0.3644
Asimetría	-0.7813	-1.2603
Curtosis	3.1142	5.9017

Paradójicamente, la proporción de valoraciones iguales a la unidad de los estados de salud es notablemente superior en nuestro caso. Pese a la aplicación de los criterios de exclusión expuestos en la sección 5.1³⁴, un 6% de las observaciones (314/4955) corresponden a valores unitarios de utilidad.

Consistencia lógica.- De los más de 3.000 pares que resultan de combinar entre sí los 78 estados valorados, un total de 558 pueden ser ordenados “lógicamente” (los niveles de uno de los estados son iguales o superiores a los del otro en cada una de las seis dimensiones). Si analizamos las utilidades medias que hemos obtenido como resultado de las valoraciones directas, únicamente en un 2.69% de los casos (15/558) se registran inconsistencias en las comparaciones dos a dos.³⁵

³⁴ Se excluyó a los sujetos que asignaron una valoración igual a la unidad (es decir, que no se mostraron dispuestos a asumir riesgo alguno de muerte en las preguntas del método PLE) a, al menos, tres de los cinco estados que les correspondió evaluar.

³⁵ En realidad, el alcance de las inconsistencias es incluso menor. En primer lugar, la primera dimensión del SF-6D, “funcionamiento físico”, no muestra una relación monótona entre niveles y gravedad del estado. En particular, los niveles 5 y 6 no tienen por qué percibirse como peores (más graves) que el nivel 4 (o incluso, que los niveles 2 y 3). Esta consideración reduciría a 14 el número de pares que revelan inconsistencias lógicas en medias. Por otra parte, algunas de las diferencias entre medias observadas no

Teniendo en cuenta que en nuestro estudio hemos optado por un diseño experimental “entre-muestras” –cada sujeto valora únicamente cinco estados y, en consecuencia, no tiene la referencia del valor asignado al resto cuando responde a las preguntas de valoración–, este resultado tiene gran interés, por cuanto confirma la posibilidad de obtener medidas de utilidad para un número suficiente de estados de salud, sin correr el riesgo de aumentar el error aleatorio en las respuestas ocasionado por la posible fatiga del entrevistado que ha de valorar un número elevado de estados.

Validez externa.- Hubiera resultado de interés testar la validez externa de las utilidades medidas directamente en nuestro estudio con el método PLE, por comparación con los valores obtenidos por Brazier et al. (2002) utilizando un método diferente (una lotería estándar “encadenada). Sin embargo, en la selección de estados de salud cuyos estadísticos descriptivos se muestran en la Tabla 4 de Brazier et al. (2002; p.280), sólo hay uno de los 78 valorados en nuestro estudio. No es posible, pues, obtener ninguna conclusión en términos de convergencia entre valoraciones, si bien puede resultar más que anecdótico el hecho de que la utilidad media del estado 423433 en nuestro estudio es 0.258 (mediana 0.26) mientras que en el de Brazier et al. es 0.58 (mediana 0.60). Este dato constituye una evidencia más de que nuestros valores son, en general, inferiores a los que sirvieron de base para la estimación del algoritmo del SF-6D, y sugiere la posibilidad de que el rango de valores obtenido por el grupo de la Universidad de Sheffield sea, al menos en parte, consecuencia del método de elicitación empleado y no, (o, al menos, no únicamente), imputable al instrumento descriptivo.

5.3. Modelización de la tarifa SF-6D

5.3.1. Modelos de efectos aleatorios

La Tabla 20 presenta los coeficientes estimados por máxima verosimilitud para dos modelos de efectos aleatorios: el modelo de efectos principales (columna 1) y una versión del modelo anterior que denominamos “modelo eficiente” (columna 2). La especificación eficiente se obtiene a partir de la de efectos principales eliminando los regresores no significativos al diez por ciento y agrupando las variables indicadoras de cualesquiera dos niveles de severidad consecutivos cuando éstos no son significativamente distintos.

resultan estadísticamente significativas, lo que implica que, en algo más de un tercio de los casos, estamos, de hecho, ante una inconsistencia “débil” (definida en términos de igualdad de medias).

Este procedimiento maximiza los grados de libertad disponibles en la estimación del modelo al tiempo que evita pequeñas inconsistencias en la predicción de la tarifa derivadas de diferencias en la estimación puntual de los coeficientes asociados a dos niveles de severidad consecutivos que no son significativamente distintos.³⁶

Tabla 20. Modelos de efectos aleatorios.

Modelo de efectos principales (1)		Modelo eficiente (2)		Modelo de efectos principales Brazier et al. (2002) (3) ^a	
FF2	-0.015			FF2	-0.058
FF3	-0.053*	FF3	-0.055*	FF3	-0.051
FF4	-0.119*	FF45	-0.117*	FF4	-0.088
FF5	-0.121*	FF6	-0.314*	FF5	-0.061
FF6	-0.313*			FF6	-0.160
LR2	0.013			LR2	-0.056
LR3	-0.041*	LR3	-0.040*	LR3	-0.076
LR4	-0.113*	LR4	-0.111*	LR4	-0.078
FS2	-0.070*	FS23	-0.072*	FS2	-0.066
FS3	-0.071*	FS4	-0.147*	FS3	-0.048
FS4	-0.146*	FS5	-0.198*	FS4	-0.066
FS5	-0.202*			FS5	-0.109
DO2	-0.004			DO2	-0.042
DO3	-0.035*	DO3	-0.029*	DO3	-0.046
DO4	-0.177*	DO45	-0.174*	DO4	-0.055
DO5	-0.175*	DO6	-0.334*	DO5	-0.103
DO6	-0.334*	SM2	-0.078*	DO6	-0.178
SM2	-0.074*	SM3	-0.094*	SM2	-0.043
SM3	-0.088*	SM4	-0.150*	SM3	-0.055
SM4	-0.147*	SM5	-0.273*	SM4	-0.115
SM5	-0.268*	VI2	-0.047*	SM5	-0.125
VI2	-0.046*	VI3	-0.107*	VI2	-0.040
VI3	-0.105*	VI4	-0.127*	VI3	-0.030
VI4	-0.128*	VI5	-0.188*	VI4	-0.040
VI5	-0.189*	EAM	0.012	VI5	-0.087
EAM	0.012	N	4955	EAM	0.074
N	4955			N	3518

Nota: ***, ** y * indican que el coeficiente es significativo al 10, 5 y 1 por ciento, respectivamente. Significatividad calculada a partir de errores estándar robustos a la presencia de heterocedasticidad (White). ^a Brazier et al. (2002) no informan del nivel de significatividad de las distintas variables, sólo de que todas son significativas al 10 por ciento.

Los coeficientes estimados en el modelo de efectos principales tienen el signo esperado, son altamente significativos, salvo tres coeficientes indicadores de nivel dos de gravedad que no son significativamente distintos de cero, y son consistentes en el

³⁶ Brazier et al. (2002) agrupan los coeficientes de cualesquiera dos variables consecutivas cuando el coeficiente estimado asociado a la indicadora de menor nivel es mayor en valor absoluto, esto es, cuando el modelo estimado produce tarifas inconsistentes. Nuestro modelo es consistente y, por tanto, nuestra preocupación es la eficiencia del mismo.

sentido de que el coeficiente asociado a un determinado nivel de gravedad nunca es significativamente inferior al correspondiente al nivel de gravedad precedente. Asimismo, destaca el valor particularmente reducido del Error Absoluto Medio (EAM), cerca de siete veces inferior al obtenido por Brazier et al. (2002) con un número de estados de salud sustancialmente mayor que el nuestro, lo que pone de manifiesto la calidad del ajuste obtenido. Los coeficientes indican que la mayor pérdida de bienestar asociada al nivel máximo de gravedad en una determinada dimensión se produce para las dimensiones “Dolor”, “Funcionamiento físico” y “Salud mental”, por orden de importancia. La misma conclusión se deriva del modelo eficiente estimado.

Un ejercicio interesante consiste en comparar los coeficientes estimados en el modelo de efectos principales con los obtenidos para el Reino Unido por Brazier et al. (2002). Para ello presentamos en columna (3) de la Tabla 20 las estimaciones puntuales obtenidas en este trabajo utilizando un modelo de efectos aleatorios con constante restringida a la unidad. Los coeficientes estimados a partir de nuestra muestra resultan ser, en la mayoría de los casos, mayores en valor absoluto a los obtenidos por Brazier et al. (2002). La pérdida de utilidad derivada de desviarse de la salud perfecta en una determinada dimensión tiende a ser mayor en nuestra muestra cuanto mayor es la desviación considerada, esto es, cuanto mayor es la gravedad considerada y es particularmente relevante en las dimensiones dolor, funcionamiento físico y salud mental.

Un precedente relevante para analizar este resultado es el trabajo de Badía et al. (2001). Estos autores comparan las valoraciones obtenidas en dos muestras representativas de las poblaciones española y británica para un mismo conjunto de estados de salud siguiendo un mismo protocolo de elicitación y utilizando el EQ-5D. La conclusión de este trabajo es que las valoraciones obtenidas no son significativamente distintas cuando las desviaciones de salud perfecta no son relevantes, esto es, la misma conclusión que parece desprenderse de la Tabla 20. No obstante, Badía et al. (2001) encuentran que los coeficientes estimados son mayores en valor absoluto (mayor pérdida) en la muestra británica cuando se trata de niveles de gravedad importantes, esto es, justo lo contrario de lo que obtenemos en nuestro estudio respecto del de Brazier et al. (2002). Esta disparidad de resultados parece indicar que las diferencias en las tarifas estimadas en muestras distintas tenderán a concentrarse en los niveles de gravedad más elevados, siendo difícilmente predecible el sentido de tal discrepancia.

Entre otras especificaciones alternativas probamos a incluir una variable indicadora de la presencia de un nivel máximo en al menos una de las dimensiones (efecto de nivel), análoga a la variable *MOST* del estudio británico, pero no resultó significativa. Asimismo, probamos una especificación en la que incluimos como regresores las variables del modelo de efectos principales y las 465 interacciones de primer orden que entre ellas pueden definirse. Treinta de estas interacciones resultaron ser significativas (28 de ellas al uno por ciento de significatividad). A cambio, tres indicadoras de efectos principales pierden su significatividad. El modelo así estimado resultó ser consistente en el sentido de no predecir utilidades superiores en estados de salud lógicamente peores a otros y su ECM fue de 0.0041, esto es, más de cuatro veces inferior al obtenido en el modelo de efectos principales. No obstante, el reducido ECM obtenido para el modelo de efectos principales unido a la sencillez del mismo nos hace decantarnos por éste modelo y, en particular, por su versión eficiente, para la obtención de una tarifa.

A continuación analizamos el papel de las características individuales en la determinación del nivel de utilidad asociado a cada estado de salud. Para ello, la Tabla 21 muestra los coeficientes estimados para un modelo de efectos principales eficiente que incluye determinadas características sociodemográficas de los encuestados como regresores adicionales. La estimación se realiza una vez eliminados quince individuos que no reportan información en algunas de las variables sociodemográficas consideradas.

Las estimaciones realizadas confirman que los encuestados valoran de forma sistemáticamente distinta un mismo estado de salud en función de su edad, su nivel de estudios, su nivel de ingresos disponible, su estado civil y del número de hijos que tienen. Por el contrario, otros regresores asociados con la actitud hacia el riesgo del encuestado como ser o no fumador habitual o el número de cigarrillos que fuma en caso de serlo no resultaron ser significativos.

El hecho de que los coeficientes asociados a las variables indicadoras de nivel de gravedad se mantengan prácticamente inalterados respecto a los obtenidos en la Tabla 20, unido a la considerable reducción en el EAM señalan que las características sociodemográficas aportan información relevante no disponible previamente en el modelo.

Tabla 21. Modelos de efectos aleatorios con características individuales

Modelo de efectos aleatorios		Modelo eficiente	
FF2	-0.021**		
FF3	-0.056*	FF3	-0.045*
FF4	-0.129*	FF45	-0.118*
FF5	-0.128*	FF6	-0.312*
FF6	-0.321*		
LR2	0.005	LR3	-0.052*
LR3	-0.047*	LR4	-0.124*
LR4	-0.119*	FS23	-0.077*
FS2	-0.076*	FS4	-0.074*
FS3	-0.072*	FS5	-0.149*
FS4	-0.149*	DO2	-0.210*
FS5	-0.211*	DO3	-0.037*
DO2	-0.008	DO45	-0.183*
DO3	-0.040*	DO6	-0.336*
DO4	-0.191*	SM2	-0.082*
DO5	-0.180*	SM3	-0.097*
DO6	-0.339*	SM4	-0.151*
SM2	-0.082*	SM5	-0.273*
SM3	-0.097*	VI2	-0.053*
SM4	-0.150*	VI3	-0.115*
SM5	-0.271*	VI4	-0.131*
VI2	-0.051*	VI5	-0.197*
VI3	-0.110*	Mujer	0.014
VI4	-0.130*	Entre 30 y 45 años	-0.048**
VI5	-0.196*	Más de 45 años	-0.041***
Mujer	0.015	Casado / convive	0.092*
Entre 30 y 45 años	-0.048**	Separado (Divorciado)/Viudo	0.023
Más de 45 años	-0.039***	Número de hijos	-0.012**
Casado / convive	0.093*	Estudios básicos	-0.026
Separado (Divorciado)/Viudo	0.026	Estudios medios	-0.033***
Número de hijos	-0.013**	Estudios superiores	-0.049**
Estudios básicos	-0.022	Ingresos bajos	0.026*
Estudios medios	-0.027***	Ingresos medio-bajos	0.052*
Estudios superiores	-0.043**	Ingresos medio-altos	0.076*
Ingresos bajos	0.025*	Ingresos altos	0.074*
Ingresos medio-bajos	0.051*	Fuma menos de 10 al día	-0.007
Ingresos medio-altos	0.074*	Fuma entre 10 y 20 al día	0.027
Ingresos altos	0.072*	Fuma entre 20 y 30 al día	-0.048
Fuma menos de 10 al día	-0.007	Fuma más de 30 al día	0.119
Fuma entre 10 y 20 al día	0.028	Tiene seguro médico privado	-0.014
Fuma entre 20 y 30 al día	-0.047	EAM	0.011
Fuma más de 30 al día	0.115	N	4940
Tiene seguro médico privado	-0.014		
EAM	0.011		
N	4940		

Nota: ***, ** y * indican que el coeficiente es significativo al 10, 5 y 1 por ciento, respectivamente. Significatividad calculada a partir de errores estándar robustos a la presencia de heterocedasticidad (White).

La inclusión de características individuales en el modelo sugiere la conveniencia de aplicar el test de Hausman para saber si el modelo de efectos aleatorios sigue siendo más eficiente que el de efectos fijos. Ese será el caso cuando no haya correlación significativa entre las características sociodemográficas incluidas y el componente individual inobservable. Los coeficientes estimados resultaron ser invariantes al procedimiento de estimación elegido (efectos fijos o aleatorios). El test de Hausman confirmó que a pesar de la inclusión de información sociodemográfica de los encuestados, el modelo de efectos aleatorios seguía siendo más eficiente que el correspondiente modelo de efectos fijos.³⁷

5.3.2. Estimaciones semiparamétricas

Una vez determinadas las características sociodemográficas relevantes a la valoración realizada por los encuestados, implementamos la primera etapa del procedimiento descrito en Abadie (2005) utilizando estas variables para dar contenido a la matriz X . El objetivo de este procedimiento es garantizar la ausencia de efectos composición en características sociodemográficas relevantes en la identificación de los coeficientes de interés. Para garantizar la comparabilidad de los coeficientes estimados utilizamos una misma distribución de características sociodemográficas de referencia. En concreto, utilizamos la distribución que las características relevantes identificadas tienen en una muestra representativa del conjunto de la población española, el Panel de Hogares de la Unión Europea, para el último año disponible, el año 2001.

Los coeficientes estimados bajo este procedimiento se presentan en la primera columna de la Tabla 22. A partir de esta especificación identificamos otra especificación, más eficiente, que presentamos en la columna (2) y que obtenemos unificando cualesquiera dos niveles de gravedad consecutivos de una misma dimensión para los que no podemos rechazar que los coeficientes estimados son iguales. Esta forma de proceder evita inconsistencias derivadas de diferencias en estimaciones puntuales que no son estadísticamente significativas.

Para llevar a cabo este contraste de forma semiparamétrica utilizamos un estimador diferencias-en-diferencias que, manteniendo una misma distribución de características sociodemográficas relevantes, analiza la significatividad de la diferencia de coeficientes

³⁷ El resultado de la estimación de efectos fijos está disponible previa petición a los autores. Obviamos su presentación en este informe para centrarnos en los resultados principales.

estimados³⁸. Al rechazar la significatividad de la diferencia de coeficientes, estimamos un nuevo coeficiente agrupando los niveles correspondientes.

Tabla 22. Modelos semiparamétricos.

Modelo de efectos aleatorios (1)		Modelo eficiente (2)	
FF2	-0.058***	FF2	-0.058***
FF3	-0.085*	FF34	-0.101*
FF4	-0.078*	FF5	-0.109*
FF5	-0.109*	FF6	-0.177*
FF6	-0.177*	LR2	-0.073**
LR2	-0.073**	LR3	-0.096*
LR3	-0.096*	LR4	-0.199*
LR4	-0.199*	FS23	-0.085*
FS2	-0.046***	FS4	-0.131*
FS3	-0.075**	FS5	-0.153*
FS4	-0.131*	DO2	-0.044***
FS5	-0.153*	DO3	-0.074*
DO2	-0.044***	DO45	-0.140*
DO3	-0.074*	DO6	-0.186*
DO4	-0.093*	SM2	-0.073**
DO5	-0.097*	SM3	-0.085*
DO6	-0.186*	SM4	-0.121*
SM2	-0.073**	SM5	-0.203*
SM3	-0.085*	VI2	-0.056***
SM4	-0.121*	VI3	-0.098*
SM5	-0.203*	VI4	-0.124*
VI2	-0.056***	VI5	-0.157*
VI3	-0.098*	EAM	0.091
VI4	-0.124*	N	4940
VI5	-0.157*		
EAM	0.091		
N	4940		

Nota: ***, ** y * indican que el coeficiente es significativo al 10, 5 y 1 por ciento de significatividad, respectivamente. Significatividad calculada a partir de errores estándar asintóticos derivados en Méndez (2008).

La comparación de los coeficientes presentados en la Tabla 22 con los obtenidos en la Tabla 20 para los correspondientes modelos de efectos aleatorios pone de manifiesto la importancia de controlar por efectos composición a la hora de obtener una tarifa. La pérdida de utilidad estimada para una desviación de la salud perfecta es, en la mayoría de los casos, inferior cuando se estima utilizando el procedimiento semiparamétrico. La

³⁸ Véase Méndez (2008) para una explicación detallada de este procedimiento.

magnitud de la diferencia entre la estimación puntual por efectos aleatorios y la estimación por el procedimiento semiparamétrico tiende a ser mayor cuanto mayor es el nivel de gravedad en cada dimensión.

Nótese que no puede predecirse a priori el sentido en el que cambiará cada coeficiente estimado en efectos aleatorios cuando la distribución de características relevantes entre los colectivos comparados sea corregida, si ha lugar, mediante el procedimiento bietápico que aplicamos. Este cambio dependerá de si la/s característica/s relevante/s sobrerrepresentada/s o infrarepresentada/s en la muestra de “tratados” tiene/n un efecto positivo o negativo en la valoración que hacen los encuestados de los estados de salud.

La discrepancia observada entre las estimaciones contenidas en las tablas 20 y 22 hace surgir de manera natural la siguiente cuestión: ¿qué estimación es más adecuada para predecir la tarifa? El hecho de que la aproximación semiparamétrica no requiera de la imposición de supuestos distribucionales en la variable dependiente, así como la flexibilidad que posee este procedimiento para la corrección de sesgos composicionales, lo convierten en firme candidato al puesto de estimador más adecuado.

El principal reparo que cabría hacerle, es que el EAM del modelo semiparamétrico multiplica por siete al correspondiente del modelo de efectos aleatorios. No obstante, éste es un resultado normal si tenemos en cuenta que el estimador de efectos aleatorios identifica los coeficientes de interés minimizando las desviaciones entre los valores observado y ajustado de la variable dependiente. Por el contrario, el estimador semiparamétrico tiene una función objetivo diferente que lo que busca es eliminar las diferencias existentes en los dos colectivos comparados en la proporción de cada valor de cada característica sociodemográfica. Es más, a pesar de tener un EAM superior al obtenido mediante el modelo de efectos aleatorios, el EAM del modelo semiparamétrico es tan solo ligeramente superior al obtenido por Brazier et al. (2002) utilizando un número de estados de salud sustancialmente superior. Por otra parte, en el modelo semiparamétrico resultan ser significativos todos los coeficientes. En particular, pasan a ser significativos los correspondientes a las variables de nivel 2 (FF2, LR2, DO2) que no lo eran en la estimación mediante efectos aleatorios.

La consideración de todos estos argumentos nos hace decantarnos por el modelo de la columna (2) de la Tabla 22 como el más apropiado para generar la tarifa.

6. Conclusiones.

6.1. Recapitulación de resultados

Resumiremos los principales resultados del estudio respetando el orden seguido en la Introducción de esta memoria. Así, en primer lugar haremos referencia a la parte más cualitativa del trabajo, la que tiene que ver con el cuestionario SF-36. Seguidamente recordaremos los resultados derivados de la medición directa de preferencias que realizamos, empleando para ello el instrumento SF-6D. Finalizaremos esta exposición comentando los resultados de las modelizaciones abordadas, subrayando las diferencias con el estudio básico de referencia de Brazier et al. (2002).

6.1.1. SF-36

El estudio realizado ha permitido recabar las respuestas de 1080 ciudadanos de la Región de Murcia a los ítems del cuestionario SF-36. Si bien las comparaciones realizadas con el conjunto nacional deben tomarse con cautela, aquellas parecen sugerir dos patrones: a) la salud de la población murciana en atributos tales como función física, rol físico y salud mental se alinea con la de la población general española; b) no es coincidente, en cambio, para el resto de dimensiones, situándose en algunos casos por encima (función social y rol emocional), y en otros casos por debajo (dolor corporal, salud general y vitalidad).

Cuando la comparación se efectúa sobre la base de puntuaciones normalizadas (tomando como norma de referencia a la población estadounidense), se verifica que las puntuaciones obtenidas en nuestro estudio rebasan a las correspondientes a la sociedad norteamericana. De nuevo hay que interpretar con prudencia este resultado, ya que la comparación en cuestión no controla por la diferente composición en edad de las dos muestras. El resto de conclusiones que podemos extraer del análisis de las respuestas dadas por la muestra murciana a las preguntas del SF-36 son, en su mayor parte, previsibles (p.ej. puntuaciones medias de la submuestra masculina que exceden de los promedios femeninos).

6.1.2. Utilidades de los estados de salud SF-6D

Una regularidad empírica bien establecida en la literatura sobre obtención de preferencias, es que los métodos “encadenados” o en dos etapas conducen a mayores utilidades que los métodos no encadenados o en una etapa. Esto se ha observado también en el ámbito de la Economía de la Salud (p.ej. Pinto y Abellán, 2005). Se han argumentado varias razones para explicar este resultado, entre otras, la transformación de la probabilidad (Bleichrodt, 1996), la aversión a las pérdidas (Oliver, 2003) e, incluso, la existencia de un posible sesgo de anclaje (Bleichrodt, 2001). Otra diferencia entre los dos tipos de procedimientos mencionados, radica en que el error aleatorio asociado a la medición puede ser mayor cuando se encadenan dos etapas consecutivas (propagación del error). En contrapartida, los procedimientos encadenados se han defendido en el contexto de la valoración de los estados de salud “leves” como una vía para dotar de mayor verosimilitud a la medición (Torrance, 1986), evitando así la aversión extrema al riesgo de los encuestados. Este último motivo es el que condujo a Brazier et al. (2002) a emplear la lotería estándar encadenada como técnica de obtención de preferencias en su estudio del SF-6D.

En contraste con la opción escogida por Brazier et al., en nuestro proyecto no se ha utilizado la lotería estándar, sino un tipo de doble lotería ya empleado previamente por algunos miembros del equipo investigador en otros estudios (p.ej. Bleichrodt et al., 2007). A priori dicho método podría conducir a utilidades más bajas que las de una lotería estándar (Bleichrodt y Schmidt, 2002). Además, se ha aplicado en este estudio de manera directa, sin encadenamiento, lo cual, a la vista de lo dicho en el párrafo anterior, también debería conducir a utilidades más reducidas.

Pues bien, tanto la media como la mediana de la distribución de utilidades directas que hemos registrado son sensiblemente más bajas que las obtenidas por Brazier et al. (2002). Más significativo aún resulta el hecho de que la utilidad media de 3 de los 78 estados que valoraron los sujetos de nuestra muestra fue negativa, por ninguna en el caso de Brazier et al. Quizá el dato más revelador de todos sea que tres cuartas partes de los encuestados otorgaron un valor inferior a -0.3 al peor estado SF-6D posible, mientras que en la muestra de Brazier et al. aproximadamente la misma proporción lo valoró por encima de cero. Esto parece indicar que, sin perjuicio de otras posibles explicaciones adicionales, la utilización de una doble lotería no encadenada ha ensanchado el rango de utilidades observadas con respecto a estudios precedentes.

6.1.3. Modelos para predecir la tarifa SF-6D

Hemos presentado dos tipos de modelos en esta memoria: el “clásico” modelo de efectos aleatorios, análogo al practicado por Brazier et al. (2002), y un nuevo modelo, basado en un enfoque semiparamétrico, más flexible que el anterior. En cualquiera de los dos casos, habida cuenta que las utilidades observadas en nuestro estudio fueron más pequeñas que las registradas por el equipo liderado por Brazier, cabría esperar que los algoritmos estimados fueran capaces de discriminar en mayor medida entre estados graves. En efecto, este es el resultado que parece hallarse, ya que los coeficientes que obtenemos en nuestras modelizaciones son mayores en valor absoluto que los del equipo británico, lo que se traduce en que el rango de puntuaciones (utilidades) que se puede predecir con nuestros modelos es más amplio que el suyo. De este modo, si con el modelo preferido por el grupo de Sheffield se obtenían utilidades comprendidas entre 0.35 y 1, con nuestro modelo de efectos aleatorios “eficiente”, el más directamente comparable con el de referencia, se generan puntuaciones entre -0.4 y 1, rompiendo así el “efecto suelo” que se ha señalado como uno de los principales inconvenientes del algoritmo. Con el modelo semiparamétrico alternativo, el rango se reduce pero continúa teniendo un recorrido sustancialmente mayor que el de Brazier et al., generando utilidades comprendidas entre -0.075 y la unidad.

Lo anterior es de gran relevancia para los fines prácticos que todo algoritmo de esta índole persigue. Se trata de poder predecir utilidades lo menos sesgadas posibles, capaces de discriminar adecuadamente entre diferentes problemas de salud. Al “descomprimirse” la zona baja de la distribución de utilidades, el SF-6D se aproxima a la sensibilidad de instrumentos como el HUI o el EQ-5D. De esta forma, podría decirse que la tarifa SF-6D derivada del presente estudio puede reflejar de forma más “equilibrada” tanto problemas de salud leves como graves.

6.2. Aplicaciones del algoritmo estimado

En los párrafos anteriores hemos inventariado los resultados más significativos que ha facilitado este proyecto. Resta ahora sugerir cuáles son las implicaciones que dichos hallazgos pueden revestir, no ya para la investigación futura, sino para su aprovechamiento práctico. A juicio del equipo investigador, el algoritmo estimado puede resultar útil al menos para dos propósitos.

Una primera utilidad es la relativa a las decisiones sanitarias meso y macro-asignativas. Tanto el SMS como la Consejería pueden, sobre la base de la tarifa, evaluar económicamente intervenciones sanitarias (tecnologías y programas). Los únicos prerequisites para hacerlo sería la necesaria identificación de los estados de salud SF-6D que mejor describen los resultados de la intervención en cuestión, poseer una estimación de las ganancias en tiempo de vida asociadas, y asimismo conocer los costes aparejados. Conocido –o al menos aproximado– todo esto, podrían calcularse ratios incrementales coste-utilidad, las cuales, por comparación con un umbral de eficiencia³⁹, permitirían saber si la tecnología evaluada “vale lo que cuesta”.

En segundo lugar, el modelo estimado brinda la oportunidad de intentar abordar un análisis coste-beneficio generalizado de la sanidad murciana. Esto es, constituye el primer paso hacia una evaluación global de la rentabilidad social de la asistencia sanitaria pública de la región. Para llevar a buen término este objetivo es necesario, sin embargo, continuar por la senda que ha abierto éste proyecto y los demás que constituyen el objeto del Convenio entre la Universidad de Murcia y la Consejería de Sanidad, realizando nuevos estudios mediante los que pueda conocerse: a) el valor social del AVAC (esto es, el valor monetario que confiere la sociedad al AVAC), y b) la contribución de la sanidad pública murciana a la mejora de la salud de los habitantes de la Región de Murcia.

³⁹ Como los identificados por el equipo investigador con ocasión del desarrollo del Proyecto 2 del Convenio: *Incorporación de la evaluación económica de tecnologías sanitarias al sistema de salud de la Región de Murcia*.

Referencias

- 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.
- Abadie A. Semiparametric difference in differences estimators. *Review of Economic Studies* 2005; 72:1-19.
- Abellán J.M. y J.L. Pinto (2000), "Quality adjusted life years as expected utilities", *Spanish Economic Review* 2, 49-63.
- Abellán-Perpiñán JM, Pinto JL, Mendez I, Badía X. Towards a better QALY model. *Health Econ* 2006, 15:665-76.
- Alonso, J., Prieto, L. y Antó, J.M. «La versión española del SF-36 Health Survey (Cuestionario de Salud SF-36): Un instrumento para la medida de los resultados clínicos». *Med Clin* 1995; 104(20), 771-6.
- Apolone, G., Cifani, S., Liberati, M.C. y Moscón, P. "Questionario sullo stato di salute SF-36. Traduzione e validazione della versione italiana: Risultati del progetto IQOLA". *Metodologia e Didattica Clinica* 1997; 5, 86-94.
- Badía X, Lizán L. "Estudios de calidad de vida". En Martín A. y J. Cano (Eds.). *Atención primaria: Conceptos, organización y práctica clínica*. Madrid: Elsevier. 2003.
- Badía X, Roset M, Herdman M, Kind P. A Comparison of United Kingdom and Spanish General Population Time Trade-off Values for EQ-5D Health States. *Med Decision Making* 2001; 21: 7-16.
- Badía X, Salamero M, Alonso J, Ollé A. La medida de salud. Guía de escalas de medición en español. Promociones y Publicaciones Universitarias. Barcelona, 1996.
- Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: Development and Final Revision of a Health Status Measure. *Med Care* 1981; 18: 787-805.

Bergner M, Bobbitt RA, Kressel S, Pollard WE, Gilson BS, Morris JR. The Sickness Impact Profile: conceptual formulation and methodological development of a health status measure. *International Journal of Health Services* 1976; 6: 393-415.

Bharmal M, Thomas J. Comparing the EQ-5D and the SF-6D descriptive systems to assess their ceiling effects in the US general population. *Value Health* 2006; 9: 262-71.

Bindman AB, Keane D, Lurie N. Measuring health changes among severely ill patients. The floor phenomenon. *Med Care* 1990;28: 1142-52.

Bleichrodt, H. (1996). Applications of utility theory in the economic evaluation of health care. Thesis Erasmus University Rotterdam, Rotterdam, The Netherlands.

Bleichrodt, H. and Schmidt, U. (2002). A context-dependent model of the gambling effect. *Management Science*, 48, 802-812.

Bleichrodt H, Johannesson M. Standard gamble, time trade-off, and rating scale: experimental results on the ranking properties of QALYs. *J Health Econ* 1997; 16: 155-75.

Bleichrodt H, Quiggin J. 'Characterizing QALYs under a general rank dependent utility model', *Journal of Risk and Uncertainty* 1997; 15: 151-65.

Brazier J, Deverill M, Green C, Harper R, Booth A. A review of the use of health status measures in economic evaluation. *Health Technology Assessment* 1999; 3(9).

Brazier J, Harper R, Jones NMB, O'Cathain A, Thomas KJ, Usherwood T, Westlake L. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *British Medical Journal* 1992; 305: 160-4.

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.

Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patients groups. *Health Econ* 2004b; 13: 873-84.

Brazier J, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care* 2004a; 42: 851-9.

Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 health survey. *J Clin Epidemiol* 1998; 51: 1115-28.

Referencias

- Brooks R et al. EuroQol: health-related quality of life measurement. Results of the Swedish questionnaire exercise. *Health Policy* 1991; 18(1): 37-48.
- Brooks R. EuroQol: the current state of play. *Health Policy* 1996; 37(1):53–72.
- Bryan S, Longworth S. Measuring health-related utility: Why the disparity between EQ-5D and SF-6D? *Eur J Health Econom* 2005; 50: 253-60.
- 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.
- Bullinger M. German translation and psychometric testing of the SF-36 Health Survey: Preliminary results from the IQOLA project, international quality of life assessment. *Soc Sci Med* 1995; 41(10): 1359-66.
- Buxton MJ, Lacey LA, Feagan BG, Niecko T, Miller DW, Townsend RJ. Mapping from disease-specific measures to utility: an analysis of the relationships between the Inflammatory Bowel Disease Questionnaire and Crohn's Disease Activity Index in Crohn's disease and measures of utility. *Value Health* 2007; 10: 214-20.
- De Bruin AF, Diederiks JPM, de Witte LP, Stevens FCJ, Philipsen H. The development of a short generic version of the Sickness Impact Profile. *J Clin Epidemiol* 1994a; 47: 407–18.
- Dolan P, Roberts J. To what extent can we explain time trade-off values from other information about respondents? *Soc Sci Med* 2002; 54: 919-29.
- Dolan P. "Modelling the relationship between description and valuation of health states". En Murray C., Salomon J., Mathers C., López A. (eds). *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications*. Geneva, World Health Organization. 2002.
- Dolan P. Modelling valuations for EuroQol health states. *Med Care* 1997; 35: 1095-108.
- Dolan P. The measurement of health-related quality of life for use in resource allocation decisions in health care. En: Culyer A J and Newhouse J P (eds.), *Handbook of health economics, vol. 1b*. Elsevier Science, Amsterdam. 2000.

Essink-Bot ML, Bonsel GJ, van der Maas PJ. Valuation of health states by the general public: Feasibility of standardised measurement procedure. *Soc Sci Med* 1990; 31: 1201-6.

Essink-Bot ML, Krabbe P, Bonsel GJ, Aaronson NK. An empirical comparison of four generic health status measures: the Nottingham Health Profile, the Medical Outcomes Study 36-item Short Form Health Survey, the COOP/WONCA charts, and the EuroQol Instrument. *Med Care* 1997; 35: 522-37.

Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zenglong Z, Depauw S, Denton M, Boyle M. Multi-attribute and single-attribute utility function for the Health Utility Index Mark 3 system. *Med Care* 2002; 40: 113-28.

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.

Garratt AM, Schmidt L, MacKintosh A, Fitzpatrick R. Quality of life measurement: Bibliographic study of patient assessed health outcome measures. *British Medical Journal* 2002; 324: 1417-21.

Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *J Clin Epidemiol*, 1993; 12: 1417-32.

Hacking H, Marcel W, Schepers V, Visser-Meily A, Lindeman E. A comparison of 3 generic health status questionnaires among stroke patients. *Journal of Stroke and Cerebrovascular Diseases* 2006; 15: 235-40.

Hawthorne G, Richardson J, Osborne R, McNeil H. The assessment of quality of life (AQoL) instrument: construction, initial validation and utility scaling. (CHPE Working Paper no 76.) Monash University, Centre for Health Program Evaluation, Melbourne. 1997.

Hornberger JC, Redelmeier DA, Petersen J. Variability among methods to assess patients well-being and consequent effect on a cost-effectiveness analysis. *J Clin Epidemiol* 1992; 45: 505-12.

Referencias

- Hunt S, McEwan P, McKenna S. *Measuring Health Status*. London: Croom Helm. 1986.
- Hunt, S., McEwen, J. and McKenna, S. Measuring health status: a new tool for clinicians and epidemiologists. *Journal of the Royal College of General Practitioners* 1985; 35: 185-88.
- Jenkinson C, Layte R, Wright L, Coulter A. *The UK SF-36: An Analysis and Interpretation Manual*. Oxford: Health Services Research Unit, University of Oxford. 1996a.
- Jenkinson C, Layte R, Wright L, Coulter A. Evidence for the sensitivity of the SF-36 health status measure to inequalities in health. *Journal of Epidemiology and Community Health* 1996b; 50: 377-80.
- Juniper EF, Norman GR, Cox FM, Roberts JN. Comparison of the standard gamble, rating scale, AOQL and SF-36 for measuring quality of life in asthma. *European Respiratory Journal* 2001; 18: 38-44.
- Katz S, Ford AB, Moskowitz RW, et al. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. *Journal of the American Medical Association* 1963; 185: 914-19.
- Kharroubi SA, Brazier J, Roberts J, O'Hagan A. Modelling SF-6D health state preference data using a nonparametric Bayesian method. *J Health Econ* 2007a; 26: 597-612.
- Kharroubi SA, Brazier J, O'Hagan A. Modelling covariates for the SF-6D standard gamble health state preference data using a nonparametric Bayesian method. *Soc Sci Med* 2007b; 64: 1242-52.
- Kharroubi SA, O'Hagan A, Brazier J. Estimating utilities from individual health preference data: a nonparametric Bayesian method. *Appl Statist* 2005; 54: 879-95.
- Kind, P. Measuring valuations for health states: Piloting the EuroQol questionnaire, Discussion paper No. 76. York: University of York, Centre for Health Economics, 1990.
- Lamers LM, Bouwmans CAM, van Straten A, Donker MCH, Hakkaart L. Comparison of EQ-5D and SF-6D utilities in mental health patients. *Health Econ* 2006; 15: 1229-36.

Lamers LM, McDonnell J, Stalmeier PFM, Krabbe PFM, Busschbach JJV. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. *Health Econ* 2006; 15: 1121-32.

Lenert LA, Cher DJ, Goldstein MK, Bergen MR, Garber AM. The effect of search procedures on utility elicitation. *Med Decision Making* 1998; 18: 76-83.

Longworth L, Bryan S. An empirical comparison of EQ-6D and SF-6D in liver transplant patients. *Health Econ* 2003; 12: 1061-67.

Maori Y, King M, Olmer L, Mozes B. A comparison of three measures: the time trade-off technique, global health-related quality of life and the SF-36 in dialysis patients. *J Clin Epidemiol* 2001; 54: 565-70.

McCabe C, Brazier J, Gilks P, Tsuchiya A, Roberts J, O'Hagan A, Stevens K. Using rank data to estimate health state utility models. *J Health Econ* 2006; 25: 418-31.

McCabe C, Stevens K, Roberts J, Brazier J. Health state values for the Health Utilities Index Mark 2 descriptive system: results from a UK valuable survey. *Health Econ* 2005; 14: 231-44.

McCallum J. The SF-36 in an Australian sample: Validating a new generic health status measure. *Australian Journal of Public Health* 1995; 19: 160-6.

McFadden D. Conditional logit analysis of qualitative Choice behaviour. En Zarembka P. (ed), *Frontiers in Econometrics*. New York Academic Press, USA. 1974: 105-42.

Méndez I. *Promoting Permanent Employment: Lessons from Spain* (manuscrito no publicado) 2008. Disponible en <http://webs.um.es/ildemm/research.html>.

Nanda U, McLendon PM, Andresen EM, Armbricht E. The SIP68: An abbreviated sickness impact profile for disability outcomes research. *Quality of Life Research* 2003; 12: 583-95.

Nelson EC, Wasson J, Kirk J, Keller A, Clark D, Dietrich A, Stewart A, Zubkoff M: Assessment of function in routine clinical practice: description of the COOP Chart method and preliminary findings. *J Chron Dis* 1987; 40 (Supl.): S55-S63.

Nelson EC, Wasson JH, Johnson DJ, Hays RD. Dartmouth COOP functional assessment charts: brief measures for clinical practice. En Spilker B. (ed) *Quality of life*

and *Pharmacoeconomics in Clinical Trials*, Second Edition. Philadelphia: Lippincott-Raven. 1996.

Nord E. EuroQol: health-related quality of life measurement. Valuations of health states by the general public in Norway. *Health Policy* 1991; 18(1): 25-36.

O'Brien BJ, Spath M, Blackhouse G, Severens JL, Dorian P, Brazier J. A view from the bridge: agreement between the SF-6D utility algorithm and the Health Utilities Index. *Health Econ* 2003; 12: 975-81.

Oliver, A. The internal consistency of the standard gamble: tests after adjusting for prospect theory. *Journal of Health Economics* 2003; 22: 659-674.

Patrick DL, Bush JW, Chen MM. Methods for measuring levels of well-being for a health status index. *Health Services Research* 1973; 8: 228-45.

Perneger TV, Leplege A, Etter JF, Rougemont A. Validation of a French language version of the MOS 36 item short form health survey (SF-36) in young healthy adults. *J Clin Epidemiol* 1995; 48: 1051-60.

Petrou S, Hockley C. An investigation into the empirical validity of the EQ-5D and SF-6D based on hypothetical preferences in a general population. *Health Econ* 2005; 14: 1169-89.

Pinto JL, Sánchez FI. *Métodos para la evaluación económica de nuevas prestaciones*. Ministerio de Sanidad y Consumo: Madrid. 2003.

Pliskin JS, Shepard DS, Weinstein MC. Utility functions for life years and health status. *Operations Research* 1980; 28: 206-23.

Read JL, Quinn DM, Berrick HV, Fineberg HV, Weinstein ML. Preferences for health outcomes: comparison of assessment methods. *Med Decision Making* 1984; 4: 315-29.

Rosser R, Kind P. A scale of valuations of states of illness: is there a social consensus? *International Journal of Epidemiology* 1978; 7: 347-358.

Sadana R. Development of standardized health state descriptions. En Murray C, Salomon J, Mathers C, López A. (eds). *Summary Measures of Population Health: Concepts, Ethics, Measurement and applications*. Geneva, World Health Organization. 2002.

Salomon JA. Reconsidering the use of rankings in the valuation of health states: a model for estimating cardinal values from ordinal data. *Population Health Metrics*

2003; 1:12. <http://www.pophealthmetrics.com/content/pdf/1478-7954-1-12.pdf> (acceso 22 de junio de 2006).

Sintonen H. The 15-D measure of HRQoL: reliability, validity and the sensitivity of its health state descriptive system. (CHPE Working Paper no 41.) Monash University, Melbourne. 1994.

Stevens K, McCabe C, Brazier J, Roberts J. Multi-attribute utility function or statistical inference models: A comparison of health state valuations models using the HUI2 health state classification system. *J Health Econom* 2007, 26: 992-1002.

Stiggelbout AM, Kiebert GM, Kievit J, Leer JWH, Stoter G, de Haes JCJM. Utility assessment in cancer patients: adjustment of time tradeoff scores for the utility of life years and comparison with standard gamble scores. *Med Decision Making* 1994; 14: 82-90.

Sullivan M, Karlsson J, Ware JE. The Swedish SF-36 Health Survey I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general population in Sweden. *Soc Sci Med* 1995; 41: 1349-58.

Szende A, Svensson K, Stahl E, Mészáros A, Berta GY. Psychometric and utility-based measures of health status of asthmatic patients with different disease control level. *Pharmacoeconomics* 2004; 22: 537-47.

Torrance GW, Feeny D, Furlong W, Barr R, Zhang Y, Wang Q. Multiattribute preference functions for a comprehensive health status classification system: health utilities index mark II. *Med Care* 1996; 34:702-2.

Torrance, G. W. Measurement of health-state utilities for economic appraisal: A review. *Journal of Health Economics* 1986; 5: 1-30.

Torrance GW. Multi-attribute utility theory as a method of measuring social preferences for health states in long-term care. En Kane RL, Kane RA (eds.): *Values in long-term care*. Lexington Books, DC Health, Lexington, MA. 1982.

Torrance GW. Social preferences for health states: an empirical evaluation of three measurement techniques. *Socio-Econ Planning Sci* 1976; 10: 129-36.

Tsuchiya A, Ikeda S, Ikegami N, Nishimura S, Sakai I, Fukuda T et al. Estimating an EQ-5D population value set: the case of Japan. *Health Econ* 2002; 11: 341-353.

Referencias

- Únala G, de Boerb J, Borsboom G, Brouwer J, Essink-Bot M, de Man R. A psychometric comparison of health-related quality of life measures in chronic liver disease. *J Clin Epidemiol* 2001; 54: 587-596.
- Van Weel C. Functional status in primary care: COOP/WONCA charts. *Disability Rehabilitation* 1993; 15: 96-101.
- Vilagut G, Ferrer M, Rajmil L, Rebollo P, Permanyer-Miralda G, Quintana JM, et al. El cuestionario de salud SF-36 español: una década de experiencia y nuevos desarrollos. *Gac Sanit* 2005; 19: 135-50.
- Ware JE Jr, Kosinski M, Dewey JE. *How to score version 2 of the SF-36 Health Survey*. Lincoln RI: Quality Metric Incorporated. 2000.
- 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.
- Ware JE Jr, Kosinski M, Turner-Bowker DM, Gandek B. *How to Score Version 2 of the SF-12 Health Survey (With a Supplement Documenting Version 1)*. Lincoln, RI: Quality Metric Incorporated. 2002.
- Ware JE, Gandek B, Keller SD, The IQoLA Project Group. Evaluating instruments used cross-nationally: Methods from the IQOLA project. En Spilker B (ed.) *Quality of Life and Pharmacoeconomics in Clinical Trials (2nd edition)*. Raven Press: New York 1996a.
- Ware JE, Kosinski M, Keller SD. A 12-item short-form health survey. Construction of scales and preliminary tests of reliability and validity. *Med Care* 1995b; 34: 220-33.
- Ware JE, Kosinski M, Keller SD. SF-12: an even shorter health survey. *Medical Outcomes Trust Bulletin* 1996; 4: 2.
- Ware JE, Kosinski M, Keller SD. *SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales, Second Edition*. Boston, Massachusetts: The Health Institute, New England Medical Center. 1995a.
- Ware JE, Kosinski M, Keller SD. *SF-36 Physical and Mental Health Summary Scales: A User's Manual*. Boston, Massachusetts: The Health Institute, New England Medical Center. 1994.

Ware JE, S  ller SD, Gandek B, Brazier JE, Sullivan M, The IQoLA Project Group. Evaluating translations of health status questionnaires: Methods from the IQOLA project". *International Journal of Technology Assessment in Health Care* 1995; 11: 525-51.

Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey 1: conceptual framework and item selection. *Med Care* 1992; 30: 473-83.

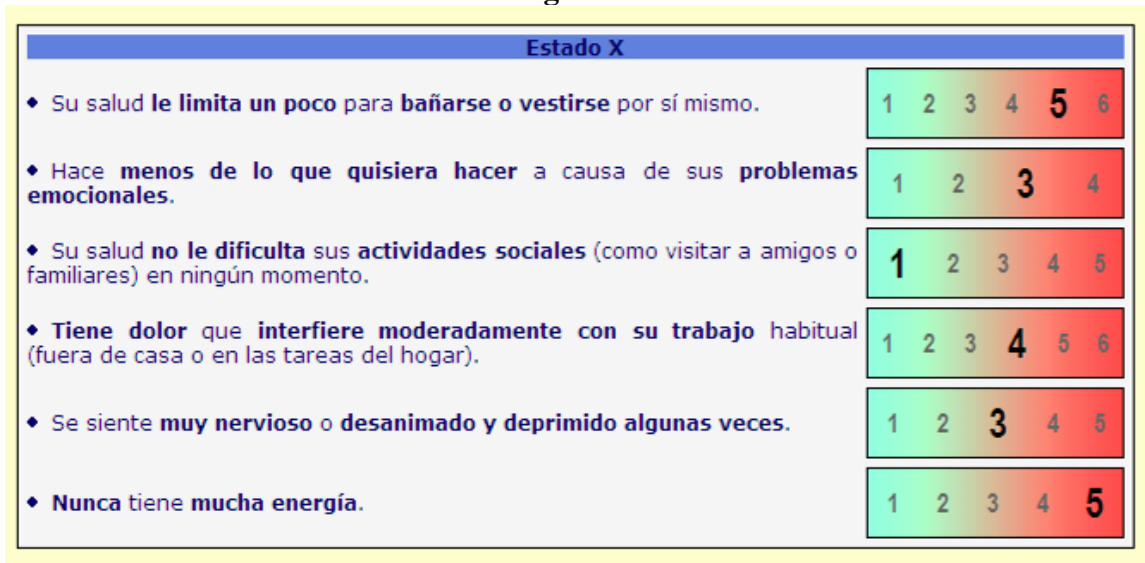
Ware JE, Snow KK, Kosinski M, Gandek B. *The SF36 Health Survey Manual and Interpretation Guide*. Boston, Massachusetts: The Health Institute, New England Medical Center. 1993.

Apéndice 1. El cuestionario

Introducción: descripción del sistema SF-6D

El cuestionario se inicia con una presentación, tras la cual se describe al entrevistado en qué consiste el sistema de clasificación de estados de salud SF-6D, explicándole los seis atributos o dimensiones que incluye, así como los niveles (entre 4 y 6) que cada dimensión puede admitir y la asociación entre niveles y “gravedad” de los problemas de salud. Con el fin de hacer más sencillo que el encuestado perciba en todos sus detalles los estados de salud que habrá de valorar a lo largo del cuestionario, se diseñaron unas “tarjetas” en las que la descripción del nivel de cada dimensión se acompañaba del número correspondiente a dicho nivel resaltado sobre un gradiente de color que iba del verde al rojo, según fuesen aumentando los problemas asociados a la dimensión en cuestión (ver Figura 1).

Figura 1



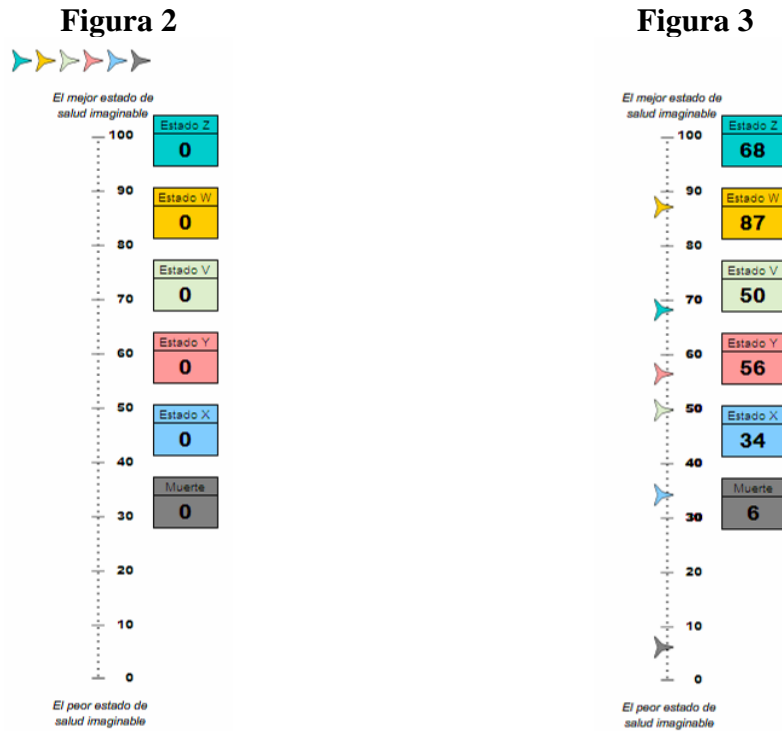
Para garantizar que el entrevistado comprendía adecuadamente esta forma de describir los estados de salud, se diseñó un “tutorial” a través del cual el estado que servía de ejemplo se iba explicando de manera desagregada dimensión por dimensión.

Puntuación de estados en la escala visual

Una vez el entrevistado ha comprendido las dimensiones y niveles que sirven para definir los estados de salud, se le presentan cinco estados de salud del sistema SF-6D, etiquetados como estados V, W, X, Y y Z (sin que la ordenación alfabética de las etiquetas pueda asociarse con el orden de “gravedad” de los estados), y se le pide que

los sitúe sobre un termómetro cuyos extremos se identifican como *el peor estado de salud imaginable* y *el mejor estado de salud imaginable*, respectivamente, y en el que también ha de situar la opción *muerte inmediata*.

En las figuras 2 y 3 se ilustra el antes y el después de esta tarea de valoración sobre la escala visual.



Obtención de medidas directas de utilidad (PLE)

La parte que contienen la aplicación del método de “doble lotería” elegido para obtener las utilidades –el *probability lottery equivalence* (PLE), que se expuso en la sección de Metodología–, se abre, para cada uno de los cinco estados valorados, con una pregunta destinada a identificar si el estado en cuestión es considerado por el sujeto “peor que la muerte”, pues el procedimiento de elecciones sucesivas que conduce al sujeto a revelar sus preferencias es diferente según se trate de un estado de salud preferido a la muerte o no. La pregunta inicial se planteó en los siguientes términos:

Figura 4

Imagine que su médico le informa de que es usted portador de una cierta enfermedad cuyo pronóstico es incierto. En concreto, en cincuenta de cada cien casos como el suyo, el paciente vive el resto de su vida sin sufrir ningún síntoma, esto es, en condiciones normales de salud. Sin embargo, en otros cincuenta casos de cada cien la enfermedad progresa dejando a los pacientes en el estado Z durante el resto de su vida.

Su médico le explica, también, que existe un tratamiento que podría curarle completamente, pero dicho tratamiento es arriesgado. En cincuenta de cada cien casos como el suyo el tratamiento es plenamente efectivo, pero en otros cincuenta casos de cada cien, provoca la muerte de manera inmediata.

En consecuencia, si usted se somete al tratamiento tendrá un 50% de probabilidad de curarse y olvidarse completamente de la enfermedad y otro 50% de probabilidad de morir. Si, por el contrario, decide no someterse

al tratamiento, se enfrenta a una probabilidad del 50% de vivir sin problemas de salud y a un 50% de probabilidad de sufrir los problemas descritos en el estado Z de manera crónica, es decir, durante el resto de su vida. Ante esta situación le pedimos, por favor, que responda:

¿Se sometería al tratamiento?



Según cuál fuese la respuesta del sujeto se le planteaba el escenario correspondiente para obtener su valoración del estado de salud mediante el método PLE. En el caso de que el sujeto eligiese la opción NO (*SIN tratamiento*), se interpretaba que prefería el estado crónico a la muerte inmediata, planteándosele a continuación el siguiente escenario:

Figura 5

Imagine ahora que se encuentra en la misma situación que le acabamos de describir, pero que el riesgo de morir asociado al tratamiento del que le hemos hablado es menor que el que antes se mencionó.

Así, si decide no someterse al tratamiento, el escenario es el mismo de la pregunta anterior: tiene usted un 50% de probabilidad de vivir sin problemas de salud y un 50% de probabilidad de sufrir los problemas descritos en el estado Z de manera crónica, es decir, durante el resto de su vida. Si se somete al tratamiento, por el contrario, las probabilidades de curarse totalmente y de morir son ahora las que se muestran en el gráfico siguiente

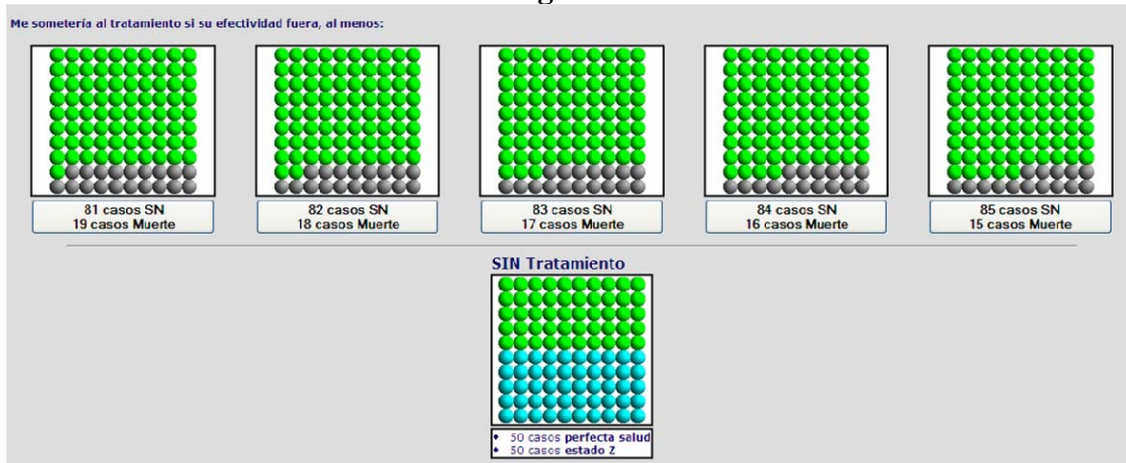
¿Se sometería al tratamiento si las alternativas fuesen las siguientes?



La probabilidad asociada a la salud normal que se ofrece como primer “estímulo” al sujeto (90% en la figura del ejemplo) es generada de manera aleatoria por la aplicación informática en el intervalo de 55 a 95. En función de cuál sea la respuesta del encuestado a esta pregunta el programa le “ofrecerá” una probabilidad menor –si elige la opción SÍ (*tratamiento*)– o una probabilidad mayor –si optó por la alternativa NO (*sin tratamiento*)–, manteniéndose constantes en el 50% las probabilidades de la salud normal y del estado crónico en la alternativa *SIN tratamiento*. Los valores que se van generando en las sucesivas preguntas siguen la secuencia basada en el procedimiento

PEST que se explica al final de este apéndice, hasta acotar un intervalo del 5% en cuyo interior la respuesta del sujeto cambia. Supongamos, por ejemplo, que el individuo elige el tratamiento para una probabilidad del 85% de disfrutar de salud normal y lo rechaza para una probabilidad de salud normal del 80%. Llegado este punto, se pide al sujeto que precise la probabilidad que haría que su decisión cambiase, planteándole una pregunta abierta como la de la Figura 6.

Figura 6



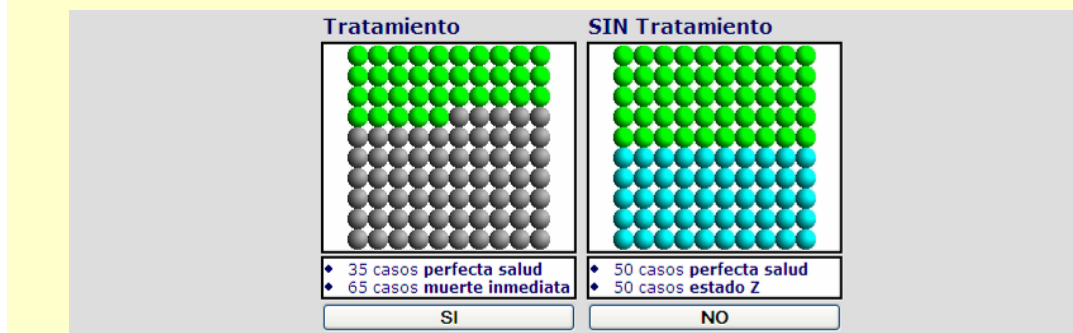
En el supuesto de que en la pregunta inicial (Figura 4) el sujeto eligiese la opción SÍ (*Tratamiento*), se interpretaba que dicho estado era considerado “peor que la muerte”⁴⁰. En tal caso, el escenario al que se enfrentaba el encuestado era como sigue:

Figura 7

Imagine ahora que se encuentra en la misma situación que le acabamos de describir, pero que el riesgo de morir asociado al tratamiento del que le hemos hablado es mayor que el que antes se mencionó.

Así, si decide no someterse al tratamiento, el escenario es el mismo de la pregunta anterior: tiene usted un 50% de probabilidad de vivir sin problemas de salud y un 50% de probabilidad de sufrir los problemas descritos en el estado Z de manera crónica, es decir, durante el resto de su vida. Si se somete al tratamiento, por el contrario, las probabilidades de curarse totalmente y de morir son ahora las que se muestran en el gráfico siguiente:

¿Se sometería al tratamiento si las alternativas fuesen las siguientes?



⁴⁰ Con el fin de evitar posibles errores, en este caso el encuestado recibía un mensaje de alerta en la pantalla en el que se le hacía ver la interpretación que se deducía de su respuesta y se le daba la opción de confirmar dicha elección o de revisarla volviendo a la pregunta inicial.

Dado que el sujeto considera el estado de salud peor que la muerte, en este caso se busca la mínima probabilidad de éxito en el tratamiento (probabilidad asociada a la salud normal en la alternativa *Tratamiento*) que el individuo exigiría con tal de no arriesgarse a sufrir el estado en cuestión con probabilidad del 50% (*SIN tratamiento*). Si el individuo elige la opción SÍ (*tratamiento*), la probabilidad asociada a salud normal en la opción *Tratamiento* (en el ejemplo, el 35%) se irá reduciendo con arreglo a la secuencia del PEST, y se irá incrementando en la medida en que el sujeto elija la opción NO (*sin tratamiento*). También en este caso el procedimiento secuencial finaliza al acotar un intervalo del 5% para, en función de esta última respuesta, terminar con la pregunta abierta en la que se pide la probabilidad exacta que consigue la indiferencia entre tratarse y no tratarse.

Autoevaluación del propio estado de salud

Uno de los objetivos del estudio era obtener información acerca de la percepción subjetiva sobre su propio estado de salud de los individuos que componían la muestra. Con esta finalidad se incluyeron tres instrumentos de medición de la salud subjetivos: el “autoclasificador” EQ-5D, el cuestionario SF-36 (versión 2) y una escala visual para que los encuestados puntuasen su estado de salud actual (comparándolo con los valores que en la misma escala asignaron a los cinco estados evaluados en la parte 1 de la encuesta).

Así, en primer lugar, el sujeto debía caracterizar su estado de salud en el momento de realizar la entrevista según las dimensiones y niveles del sistema EQ-5D (Figura 8).

Figura 8

<p>Movilidad</p> <p><input type="radio"/> No tengo problemas para caminar ■</p> <p><input type="radio"/> Tengo algunos problemas para caminar ■</p> <p><input type="radio"/> Tengo que estar en cama ■</p>	<p>Cuidado Personal</p> <p><input type="radio"/> No tengo problemas para realizar mi cuidado personal ■</p> <p><input type="radio"/> Tengo algunos problemas para lavarme o vestirme ■</p> <p><input type="radio"/> Soy incapaz de lavarme o vestirme por mí mismo ■</p>
<p>Actividades Cotidianas (ej.:trabajar, estudiar, hacer las tareas domésticas, disfrutar del tiempo libre)</p> <p><input type="radio"/> No tengo problemas para realizar mis actividades cotidianas ■</p> <p><input type="radio"/> Tengo algunos problemas para realizar mis actividades cotidianas ■</p> <p><input type="radio"/> Soy incapaz de realizar mis actividades cotidianas ■</p>	<p>Dolor, Molestias</p> <p><input type="radio"/> No tengo dolor ni malestar ■</p> <p><input type="radio"/> Tengo dolor o malestar moderado ■</p> <p><input type="radio"/> Tengo mucho dolor o mucho malestar ■</p>
<p>Ansiedad o Depresión</p> <p><input type="radio"/> No estoy ansioso ni deprimido ■</p> <p><input type="radio"/> Estoy moderadamente ansioso o deprimido ■</p> <p><input type="radio"/> Estoy muy ansioso o deprimido ■</p>	

A continuación, los entrevistados debían cumplimentar la versión 2 del cuestionario de salud SF-36, que se administró dentro de la misma aplicación informática mediante un sistema de “botones-radio” (ver el extracto de la Figura 9, a modo de ejemplo). Esta

parte de la aplicación (también el autoclasificador EQ-5D) se programó de manera que resultase imposible avanzar en la encuesta si se había dejado sin responder alguna de las preguntas (o de los ítems que componen cada una de ellas). De este modo se garantizó un 100% de respuestas válidas en esta parte de la encuesta.

Figura 9

4. Durante las 4 últimas semanas, ¿con qué frecuencia ha tenido alguno de los siguientes problemas en su trabajo o en sus actividades cotidianas, a causa de su salud física?

	Siempre	Casi siempre	Algunas veces	Sólo alguna vez	Nunca
¿Tuvo que reducir el tiempo dedicado al trabajo o a sus actividades cotidianas?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
¿Hizo menos de lo que hubiera querido hacer?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
¿Tuvo que dejar de hacer algunas tareas en su trabajo o en sus actividades cotidianas?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
¿Tuvo dificultad para hacer su trabajo o sus actividades cotidianas (por ejemplo, le costó más de lo normal)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5. Durante las 4 últimas semanas, ¿con qué frecuencia ha tenido alguno de los siguientes problemas en su trabajo o en sus actividades cotidianas, a causa de algún problema emocional (como estar triste, deprimido o nervioso)?

	Siempre	Casi siempre	Algunas veces	Sólo alguna vez	Nunca
¿Tuvo que reducir el tiempo dedicado al trabajo o a sus actividades cotidianas por algún problema emocional ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
¿Hizo menos de lo que hubiera querido hacer por algún problema emocional ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
¿Tuvo su trabajo o sus actividades cotidianas menos cuidadosamente que de costumbre, por algún problema emocional ?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. Durante las 4 últimas semanas, ¿hasta qué punto su salud física o los problemas emocionales han dificultado sus actividades sociales habituales con la familia, los amigos, los vecinos u otras personas?

	Nada	Un poco	Regular	Dastante	Mucho
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. ¿Tuvo dolor en alguna parte del cuerpo durante las 4 últimas semanas?

	No, ninguno	Sí, muy poco	Sí, un poco	Sí, moderado	Sí, mucho	Sí, muchísimo
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Por último, se pedía a los entrevistados que valorasen su propio estado de salud sobre la misma escala que al inicio del cuestionario sirvió para valorar los cinco estados de salud evaluados y la muerte. La escala con las valoraciones de la parte 1 de la encuesta (ver Figura 3) era “recuperada” a tal fin por la aplicación informática en esta última parte del cuestionario.

Características socio-demográficas

La última parte del cuestionario se dedicó a recoger información sobre características del entrevistado que posteriormente habría de utilizarse en el proceso de análisis de los datos. El formulario se reproduce a continuación:

Figura 10

Edad:

Sexo:

Estado Civil:

Nº de hijos que conviven con usted:

Nº de personas, incluido usted mismo, que conviven en el hogar:

Nivel de estudios terminados:

Nivel de ingresos mensuales del hogar:

Número de cigarrillos que usted fuma al día:

¿Tiene usted alguna póliza de seguro médico privada?

Finalmente se pedía a los entrevistados que indicasen el grado de dificultad que les había supuesto responder a las diferentes preguntas contenidas en la encuesta.

Procedimiento para la obtención de los valores de indiferencia (PEST) ⁴¹

En la parte de la encuesta en la que se obtienen las valoraciones de los estados de salud mediante el procedimiento PLE, el valor de indiferencia se identificó siguiendo un sistema de obtención de respuestas secuencial como el sugerido por Luce (2000), con el fin de evitar los sesgos asociados a los métodos basados en *matching*. El valor del parámetro con el que se pretende conseguir la indiferencia –en nuestro caso, la probabilidad asociada al resultado “salud normal” en la alternativa *Tratamiento*– va cambiando según la secuencia que más adelante se describe, en función de las elecciones que realiza el sujeto, hasta acotar el mínimo intervalo del 5%; llegados a este punto se pide al entrevistado que concrete (con precisión de 1 punto porcentual) la probabilidad que le hace ser indiferente entre recibir tratamiento o no.

Las normas que siguió el procedimiento en nuestro estudio fueron las siguientes:

1. El valor inicialmente propuesto se determina de manera aleatoria dentro del rango de valores posibles (de 55% a 95% para estados “mejores que la muerte”, de 5% a 45% para estados “peores que la muerte”).
2. Tras la primera respuesta del sujeto, la magnitud de la duración que se ofrece al sujeto aumenta (disminuye) en 10 puntos porcentuales (“paso inicial”).
3. Una segunda elección coincidente –por ejemplo, elige “NO (*Sin tratamiento*)” dos veces seguidas– da lugar al mismo incremento (disminución) en la probabilidad ofrecida.
4. A partir de la cuarta elección en un mismo sentido, el incremento (disminución) se duplica, siempre y cuando no se excedan los límites; en caso contrario, se fija una nueva probabilidad con un valor distante en una unidad de dicho límite.
5. Cada vez que el sujeto cambia su elección (pasa, por ejemplo, de elegir “SÍ (*Tratamiento*)” a elegir “NO (*Sin Tratamiento*)”, el incremento (disminución) en la probabilidad se divide por dos.
6. La tercera elección en el mismo sentido puede dar lugar a un incremento (disminución) en la probabilidad ofrecida de igual o doble magnitud que el

⁴¹ Acrónimo de *Parameter Estimation by Sequential Testing*.

anterior, según cuál haya sido el último cambio de elección. Si éste fue precedido de un aumento en la magnitud del incremento (disminución), ésta se mantiene; en caso contrario la magnitud de la variación se duplica.

7. Cada dos elecciones se ofrece al sujeto una distribución de probabilidades en la alternativa “*SÍ (Tratamiento)*” totalmente aleatoria, sin relación alguna con la secuencia utilizada. Con ello se intenta que el sujeto no perciba que se le está “conduciendo” a lo largo de un proceso de convergencia.

Apéndice 2. Resultados del SF-36

Descripción de la muestra

Los resultados que se van a mostrar en esta sección corresponden a la totalidad de las 1.080 encuestas realizadas. Así, se incluyen tanto las 1.020 encuestas dirigidas a la obtención de la tarifa española para el SF-6D, como las otras 60 encuestas destinadas a validar el procedimiento de obtención de las utilidades. Podemos entender la muestra total como el resultado de la agregación de 18 grupos (17 grupos más 1) de 60 observaciones cada uno. De hecho, el diseño muestral se realizó a partir de la composición de cada uno de esos grupos, con la finalidad de que se ajustara a las cuotas de edad y sexo de la población española.

Tabla 23. Distribución de las observaciones por edad y sexo, dentro de cada grupo.

	Hombres	Mujeres
Entre 18 y 29 años.	7	6
Entre 30 y 41 años.	8	8
Entre 42 y 53 años.	7	7
Entre 54 y 65 años.	5	5
66 y más años.	3	4
Total	30	30

Como resulta frecuente en los procesos de muestreo se produjeron pequeños desajustes en alguno de los grupos de encuestación, si bien el resultado final de la agregación no difiere significativamente del comportamiento por grupos de edad y sexo buscado.

Tabla 24. Porcentajes por grupo de edad y sexo para el total de la muestra.

	Hombres	Mujeres	Total
Entre 18 y 29 años.	11,85	10,09	21,94
Entre 30 y 41 años.	13,15	13,43	26,57
Entre 42 y 53 años.	11,57	11,48	23,06
Entre 54 y 65 años.	8,43	8,33	16,76
66 y más años.	5,00	6,67	11,67
Total (1.080 obs)	50,00	50,00	100,00

La edad media en la muestra es de poco más de 43 años, y la mediana se encuentra situada en 42 años. Para el caso de la submuestra masculina, la media se situó en poco más de 42 años y medio, y la mediana en 41 años y medio. En la submuestra femenina, la media toma un valor de algo más de 44 años, y la mediana se encuentra en 43 años.

Con independencia del diseño muestral, y a efectos de análisis, se ha utilizado una división por grupos de edad distinta. La razón última de esta elección se encuentra en la mayor riqueza informativa fruto del mayor grado de desagregación. A continuación se

muestran las observaciones correspondientes a cada uno de los nuevos grupos de edad por sexos.

Tabla 25. Distribución de las observaciones por grupos de edad y sexo.

	Hombres	Mujeres	Total
Menor que 25	90	81	171
Entre 25 y 34 años	126	115	241
Entre 35 y 44 años	80	83	163
Entre 45 y 54 años	107	113	220
Entre 55 y 64 años	76	74	150
Entre 65 y 74 años	42	45	87
75 y más años	19	29	48
Total	540	540	1.080

Como se puede observar, aproximadamente una sexta parte de la muestra está formada por individuos menores de 25 años. El grupo más numeroso es el formado por los individuos entre 25 y 34 años, que supone, aproximadamente, un 22% del total de observaciones. El grupo de edad intermedia, comprendido entre 35 y 44 años, y 45 y 54 años, supone, un sexto y un quinto de la muestra, respectivamente. Finalmente, los tres grupos de mayor edad, entre 55 y 64 años, entre 65 y 74 años, y mayores de 75 años, representan, respectivamente, un 13%, 8% y 4,5%.

También, y a efectos de análisis, se considerarán diversas categorías atendiendo a los niveles de estudios terminados, el nivel de ingresos mensual del hogar, el estado civil del encuestado, la presencia o no del hábito tabáquico y si dispone o no de un seguro médico privado. Los resultados de aplicar estas taxonomías a la muestra total, quedan recogidos en la siguiente tabla.

Tabla 26. Porcentajes en la muestra según niveles de estudios, de ingresos, estado civil, tabaquismo y seguro sanitario.

Niveles de Estudios	%	Estado Civil	%
Sin estudios	10,37	Soltero	34,07
Primarios	23,7	Casado o con pareja de hecho	59,72
Secundarios	35,28	Separado, Viudo, Divorciado	6,2
Superiores	30,65		
Nivel de Ingresos	%		%
Hasta 900 euros	7,22	No Fumador	72,13
Entre 901 y 1.500 euros	16,57	Fumador	27,87
Entre 1.501 y 2.000 euros	28,52		%
Entre 2.001 y 3.000 euros	29,54	Sin seguro sanitario privado	80,56
Más de 3.000 euros	18,15	Seguro sanitario privado	19,44

Como se puede observar, poco más de un 10% de la muestra no alcanzó a terminar los estudios primarios, casi un cuarto de la muestra tiene estudios primarios, y aproximadamente dos terceras partes de la muestra concluyó sus estudios secundarios o superiores. Por otro lado, y en relación a los niveles de ingresos mensuales del hogar, se obtuvo que poco más de un 7% de los encuestados declaran unos ingresos mensuales inferiores a los 900 euros, aproximadamente una sexta parte de la muestra señala recibir una renta comprendida entre dicha cantidad y 1500 euros; casi un 60% de la muestra expresa disfrutar de unos ingresos medios mensuales en su hogar comprendidos entre 1500 y 3000 euros, y, finalmente, algo menos de un 20% de la muestra asegura recibir una renta superior a dicho límite.

En lo concerniente al estado civil, poco más de una tercera parte de la muestra declara ser soltera, casi un 60% asegura estar casado o convivir maritalmente con otra persona, y poco más de 6% señala encontrarse en otra situación (viudo/a; separado/a y divorciado/a). En lo que se refiere al tabaquismo, podemos señalar que poco más de una cuarta parte de la muestra se declara fumador, y el resto, tres cuartas partes, asegura ser no fumador. Por último, y en lo referido a si el encuestado disfruta de algún tipo de seguro médico privado, aproximadamente cuatro quintas partes de nuestra muestra señala no contar con ningún seguro de este tipo, mientras que casi uno de cada cinco encuestados reconoce gozar de dicha cobertura.

Análisis de las puntuaciones SF-36

En las siguientes tablas, se muestran los resultados en términos de puntuaciones en las escalas del SF-36 normalizadas y sin normalizar. Además de los valores para la totalidad de la muestra, se han calculado valores por submuestras, atendiendo al sexo, grupos de edad, nivel de estudios, nivel de ingresos, la tenencia o no de un seguro médico privado y la relación con el tabaco del encuestado. Todos los valores que se muestran se han calculado sobre la muestra bruta, esto es, no se ha realizado ningún tipo de ajuste por submuestras para asegurar la eliminación de efectos composicionales, como los derivados de la edad.⁴²

⁴² En el ámbito epidemiológico es frecuente la utilización de comparaciones estandarizadas, de forma directa o indirecta, con el fin de asegurar la eliminación de efectos composicionales. En este caso, el efecto a eliminar sería el relacionado con la edad. Dado que existe una relación negativa entre la edad del encuestado y el estado de salud disfrutado, puede tener sentido en las comparaciones ajustar previamente por la edad.

No existe una norma claramente identificada en lo que se refiere al comportamiento de la población española de acuerdo a las puntuaciones del SF-36. Por tanto, resulta relativamente complejo juzgar los valores aquí obtenidos. Probablemente, la mejor de entre las opciones existentes es la que proporcionan en su artículo Vilagut et al (2005). En dicho trabajo, proporcionan unos intervalos de confianza al 95% para las puntuaciones de las distintas escalas, que obtuvieron a partir de una estimación conjunto de una serie, relativamente amplia, de trabajos anteriores sobre población española, tanto general, como afectada por distintas dolencias (p.ej. EPOC: asma, VIH, etc). Si realizamos la comparación con el total de la muestra, podremos observar como los valores obtenidos en la dimensión Función Física, se encuentran dentro del intervalo obtenido en ese trabajo. Lo mismo ocurre para la puntuación de la dimensión Rol Físico y Salud Mental. Por otro lado, las puntuaciones obtenidas por la muestra murciana para las dimensiones Dolor Corporal, Salud General y Vitalidad, se sitúan por debajo del extremo inferior del intervalo de confianza (en los dos primeros casos, casi 10 puntos básicos). Mientras, para las dimensiones Función Social y Rol Emocional las valoraciones de nuestra muestra se encuentra por encima del extremos superior del intervalo.

Una segunda posibilidad de comparación es la proporcionada por el trabajo de Alonso et al (1998) que, hasta donde sabemos, es el único trabajo de gran tamaño muestral (más de 9.000 observaciones) sobre población general española. Las conclusiones de esa comparación, serían muy similares a las enunciadas anteriormente.

En lo que se refiere a las comparaciones intra-muestra, se obtienen resultados previsibles. En primer lugar, se puede observar como las puntuaciones medias de la submuestra masculina exceden a las de la femenina. Este resultado se encuentra habitualmente en las encuestadas destinadas a comprobar el estado de salud de la población. En la misma dirección, y como era esperable, también se observa, tanto para hombres, como para mujeres, un claro deterioro de las puntuaciones obtenidas en todas las dimensiones conforme se incrementa la edad del grupo. En lo que concierne a la relación entre el nivel de estudios terminados y el estado de salud, parece obtenerse evidencia de que existe una relación positiva entre la puntuación obtenida y el mayor nivel de estudios. Al menos, existe un claro comportamiento diferencia en lo que se refiere a tener al menos estudios secundarios, con respecto a menores niveles de estudios terminados. Este fenómeno se puede explicar de una doble perspectiva. De un

lado, el nivel de estudios terminados se puede ver como una variable Proxy de la clase social. De otro lado, existe un claro componente composicional de la edad, dado que en España, los menores niveles de estudios terminados se concentran, especialmente, en los grupos de mayor edad. Estos mismos dos factores se pueden utilizar con la finalidad de explicar algunas diferencias en términos de puntuación que se registran entre los distintos subgrupos en términos de renta, en especial, los valores más bajos del grupo que disfruta de menor nivel de renta. Sea como fuere, y como suele ser habitual, las rentas declaradas por los encuestados deben ser tomadas con bastante precaución, puesto que puede existir un importante bolsa de respuestas no acordes con la realidad. En lo que se refiere al estado de salud, podemos ver como las mejores puntuaciones se concentran, como era de esperar, en el grupo de los solteros. Este fenómeno, claramente, se explica por la distinta composición en términos de edad de los grupos. En lo que respecta a la tenencia de un seguro médico privado, no se advierte que exista un comportamiento diferencial claro. Finalmente, y de forma contraintuitiva, se observa un resultado positivo en términos de diferencias en las puntuaciones entre fumadores y no fumadores. Detrás de este resultado inesperado, puede existir un efecto composicional en las edades de ambos grupos.

Finalmente, como se apuntó, se presenta los valores normalizados utilizando la población norteamericana como referencia. Se puede observar, como tanto para la totalidad de la muestra, como por sexos, los valores obtenidos en la muestra murciana son superiores a los de la muestra norteamericana en todas las escalas salvo en la referida a la salud general, que se encuentra ligeramente por debajo del valor poblacional. De nuevo, habría que tomar este resultado con precauciones, puesto que estamos comparando dos grupos heterogéneos en cuanto a su composición por edades.

Tabla 27. Puntuación en las escalas del SF-36.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	90,60	88,83	76,26	66,67	70,19	90,71	93,33	79,44
Observaciones	1080	1080	1080	1080	1080	1080	1080	1080
Percentil 5°	55,00	50,00	40,00	30,00	40,63	50,00	58,33	50,00
Percentil 25°	90,00	81,25	60,00	55,00	62,50	87,50	100,00	72,50
Mediana	95,00	100,00	80,00	70,00	75,00	100,00	100,00	80,00
Percentil 75°	100,00	100,00	100,00	80,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	95,00	100,00	100,00	100,00	100,00
Desv. Típica	15,82	19,03	21,93	19,22	16,75	16,84	14,88	15,31
Mínimo	0,00	0,00	0,00	5,00	0,00	0,00	0,00	10,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 28. Puntuación en las escalas del SF-36, submuestra masculina.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	92,25	90,10	78,31	67,98	72,22	92,59	94,86	81,99
Observaciones	540	540	540	540	540	540	540	540
Percentil 5°	65,00	50,00	40,00	30,00	43,75	62,50	66,67	57,50
Percentil 25°	90,00	87,50	70,00	55,00	62,50	87,50	100,00	75,00
Mediana	100,00	100,00	80,00	70,00	75,00	100,00	100,00	85,00
Percentil 75°	100,00	100,00	100,00	80,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Desv. Típica	14,88	19,34	21,34	19,50	16,07	15,55	13,63	13,44
Mínimo	0,00	0,00	0,00	5,00	6,25	0,00	0,00	15,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 29. Puntuación en las escalas del SF-36, submuestra femenina.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	88,95	87,55	74,20	65,36	68,16	88,82	91,79	76,90
Observaciones	540	540	540	540	540	540	540	540
Percentil 5°	50,00	50,00	30,00	30,00	37,50	50,00	50,00	45,00
Percentil 25°	85,00	75,00	60,00	55,00	56,25	75,00	91,67	70,00
Mediana	95,00	100,00	80,00	70,00	68,75	100,00	100,00	80,00
Percentil 75°	100,00	100,00	100,00	80,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	95,00	93,75	100,00	100,00	100,00
Desv. Típica	16,56	18,65	22,32	18,86	17,18	17,85	15,91	16,59
Mínimo	10,00	0,00	0,00	10,00	0,00	0,00	25,00	10,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 30. Puntuación en las escalas del SF-36, submuestra menores de 25 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	97,49	94,12	82,28	75,94	75,04	92,18	92,74	81,17
Observaciones	171	171	171	171	171	171	171	171
Percentil 5°	90,00	68,75	40,00	45,00	43,75	62,50	58,33	55,00
Percentil 25°	100,00	100,00	70,00	65,00	62,50	87,50	91,67	75,00
Mediana	100,00	100,00	80,00	75,00	75,00	100,00	100,00	85,00
Percentil 75°	100,00	100,00	100,00	90,00	87,50	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Desv. Típica	8,51	15,89	21,42	17,07	17,54	15,58	15,98	15,45
Mínimo	0,00	0,00	0,00	25,00	25,00	25,00	0,00	15,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 31. Puntuación en las escalas del SF-36, submuestra entre 25 y 34 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	97,28	94,19	83,53	72,30	74,59	95,23	95,02	80,98
Observaciones	241	241	241	241	241	241	241	241
Percentil 5°	90,00	75,00	50,00	40,00	50,00	75,00	75,00	60,00
Percentil 25°	95,00	93,75	70,00	65,00	62,50	100,00	91,67	75,00
Mediana	100,00	100,00	80,00	75,00	75,00	100,00	100,00	85,00
Percentil 75°	100,00	100,00	100,00	85,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Desv. Típica	5,84	13,31	18,45	17,51	13,67	10,42	11,33	12,09
Mínimo	60,00	0,00	0,00	15,00	31,25	37,50	0,00	35,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 32. Puntuación en las escalas del SF-36, submuestra entre 35 y 44 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	94,60	91,18	80,43	72,24	71,28	91,72	94,22	79,94
Observaciones	163	163	163	163	163	163	163	163
Percentil 5°	75,00	50,00	40,00	40,00	43,75	50,00	58,33	50,00
Percentil 25°	95,00	87,50	70,00	65,00	62,50	87,50	100,00	70,00
Mediana	100,00	100,00	80,00	75,00	75,00	100,00	100,00	85,00
Percentil 75°	100,00	100,00	100,00	85,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	95,00	100,00	100,00	100,00	100,00
Desv. Típica	9,97	17,36	20,74	18,16	18,20	16,51	13,84	16,60
Mínimo	40,00	6,25	10,00	5,00	0,00	25,00	33,33	10,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 33. Puntuación en las escalas del SF-36, submuestra entre 45 y 54 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	92,39	89,74	75,68	66,02	69,63	90,80	93,52	76,98
Observaciones	220	220	220	220	220	220	220	220
Percentil 5°	75,00	50,00	40,00	40,00	50,00	50,00	58,33	45,00
Percentil 25°	90,00	81,25	60,00	55,00	62,50	87,50	95,83	70,00
Mediana	95,00	100,00	80,00	70,00	68,75	100,00	100,00	80,00
Percentil 75°	100,00	100,00	100,00	75,00	78,13	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	90,00	93,75	100,00	100,00	100,00
Desv. Típica	9,67	16,81	20,34	15,47	14,83	16,74	14,49	15,77
Mínimo	40,00	0,00	10,00	20,00	12,50	0,00	0,00	20,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 34. Puntuación en las escalas del SF-36, submuestra entre 55 y 64 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	85,50	86,08	67,87	58,43	65,67	88,25	91,06	78,73
Observaciones	150	150	150	150	150	150	150	150
Percentil 5°	50,00	50,00	20,00	25,00	31,25	37,50	50,00	45,00
Percentil 25°	85,00	75,00	60,00	50,00	50,00	75,00	91,67	70,00
Mediana	90,00	93,75	70,00	60,00	68,75	100,00	100,00	80,00
Percentil 75°	95,00	100,00	80,00	70,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	85,00	93,75	100,00	100,00	100,00
Desv. Típica	16,26	18,87	21,41	17,05	17,84	19,39	17,11	16,71
Mínimo	10,00	25,00	0,00	5,00	18,75	0,00	25,00	20,00
Máximo	100,00	100,00	100,00	95,00	100,00	100,00	100,00	100,00

Tabla 35. Puntuación en las escalas del SF-36, submuestra entre 65 y 74 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	73,62	74,50	63,56	51,44	62,43	86,21	94,25	80,80
Observaciones	87	87	87	87	87	87	87	87
Percentil 5°	35,00	37,50	20,00	25,00	37,50	50,00	66,67	45,00
Percentil 25°	60,00	50,00	50,00	35,00	50,00	75,00	100,00	75,00
Mediana	80,00	75,00	70,00	55,00	62,50	100,00	100,00	85,00
Percentil 75°	90,00	93,75	80,00	65,00	75,00	100,00	100,00	90,00
Percentil 95°	95,00	100,00	100,00	80,00	81,25	100,00	100,00	100,00
Desv. Típica	20,26	22,64	21,73	16,84	15,08	17,89	14,11	15,44
Mínimo	10,00	12,50	10,00	15,00	18,75	25,00	25,00	30,00
Máximo	100,00	100,00	100,00	85,00	100,00	100,00	100,00	100,00

Tabla 36. Puntuación en las escalas del SF-36, submuestra 75 y más años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	57,50	65,36	56,04	42,81	57,94	74,74	88,37	75,00
Observaciones	48	48	48	48	48	48	48	48
Percentil 5°	15,00	12,50	10,00	10,00	31,25	25,00	25,00	45,00
Percentil 25°	32,50	46,88	40,00	30,00	50,00	62,50	87,50	65,00
Mediana	60,00	71,88	60,00	40,00	59,38	75,00	100,00	80,00
Percentil 75°	80,00	93,75	70,00	60,00	68,75	100,00	100,00	90,00
Percentil 95°	95,00	100,00	100,00	70,00	81,25	100,00	100,00	95,00
Desv. Típica	27,17	28,94	22,48	18,85	16,14	24,93	22,66	16,18
Mínimo	0,00	0,00	10,00	5,00	12,50	12,50	25,00	30,00
Máximo	95,00	100,00	100,00	80,00	87,50	100,00	100,00	100,00

Tabla 37. Puntuación en las escalas del SF-36, submuestra masculina menores de 25 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	97,39	93,40	82,78	77,67	75,49	92,92	93,70	83,11
Observaciones	90	90	90	90	90	90	90	90
Percentil 5°	90,00	37,50	40,00	45,00	43,75	62,50	66,67	60,00
Percentil 25°	100,00	100,00	70,00	70,00	68,75	87,50	100,00	80,00
Mediana	100,00	100,00	90,00	80,00	75,00	100,00	100,00	85,00
Percentil 75°	100,00	100,00	100,00	90,00	87,50	100,00	100,00	95,00
Percentil 95°	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Desv. Típica	11,05	19,71	21,62	17,14	17,26	15,79	16,45	15,52
Mínimo	0,00	0,00	0,00	30,00	25,00	25,00	0,00	15,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 38. Puntuación en las escalas del SF-36, submuestra masculina entre 25 y 34 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	97,82	93,40	84,21	71,55	75,50	97,02	95,30	82,42
Observaciones	126	126	126	126	126	126	126	126
Percentil 5°	90,00	75,00	60,00	25,00	56,25	75,00	75,00	60,00
Percentil 25°	100,00	93,75	70,00	65,00	62,50	100,00	100,00	75,00
Mediana	100,00	100,00	80,00	75,00	75,00	100,00	100,00	85,00
Percentil 75°	100,00	100,00	100,00	85,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Desv. Típica	5,67	15,40	17,95	18,75	12,91	6,97	11,95	10,84
Mínimo	60,00	0,00	0,00	15,00	43,75	75,00	0,00	40,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 39. Puntuación en las escalas del SF-36, submuestra masculina entre 35 y 44 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	95,19	92,89	82,38	72,25	73,67	92,97	96,56	81,25
Observaciones	80	80	80	80	80	80	80	80
Percentil 5°	77,50	50,00	40,00	40,00	46,88	56,25	75,00	52,50
Percentil 25°	95,00	93,75	70,00	65,00	62,50	100,00	100,00	75,00
Mediana	100,00	100,00	80,00	75,00	75,00	100,00	100,00	85,00
Percentil 75°	100,00	100,00	100,00	85,00	84,38	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	97,50	100,00	100,00	100,00	100,00
Desv. Típica	9,46	18,11	20,39	18,69	17,86	16,26	9,62	15,27
Mínimo	40,00	6,25	10,00	5,00	6,25	25,00	50,00	15,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 40. Puntuación en las escalas del SF-36, submuestra masculina entre 45 y 54 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	93,83	90,13	78,41	67,66	73,01	92,29	94,78	81,96
Observaciones	107	107	107	107	107	107	107	107
Percentil 5°	70,00	50,00	30,00	40,00	50,00	50,00	66,67	60,00
Percentil 25°	95,00	87,50	70,00	60,00	62,50	87,50	100,00	75,00
Mediana	95,00	100,00	80,00	70,00	75,00	100,00	100,00	85,00
Percentil 75°	100,00	100,00	100,00	80,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	95,00	93,75	100,00	100,00	100,00
Desv. Típica	9,85	18,07	20,47	16,30	13,21	16,70	13,90	12,47
Mínimo	45,00	0,00	10,00	30,00	37,50	0,00	0,00	50,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 41. Puntuación en las escalas del SF-36, submuestra masculina entre 55 y 64 años

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	87,63	88,73	60,86	60,86	68,75	90,13	94,63	81,18
Observaciones	76	76	76	76	76	76	76	76
Percentil 5°	55,00	50,00	20,00	35,00	37,50	50,00	58,33	55,00
Percentil 25°	85,00	84,38	60,00	50,00	56,25	87,50	100,00	75,00
Mediana	90,00	100,00	70,00	60,00	68,75	100,00	100,00	85,00
Percentil 75°	95,00	100,00	80,00	72,50	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	90,00	100,00	100,00	100,00	100,00
Desv. Típica	13,63	18,82	22,38	18,23	18,00	17,71	13,10	14,69
Mínimo	35,00	25,00	10,00	5,00	25,00	25,00	41,67	20,00
Máximo	100,00	100,00	100,00	95,00	100,00	100,00	100,00	100,00

Tabla 42. Puntuación en las escalas del SF-36, submuestra masculina entre 65 y 74 años

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	76,19	78,87	67,86	52,62	64,43	88,69	97,22	83,33
Observaciones	42	42	42	42	42	42	42	42
Percentil 5°	30,00	43,75	40,00	30,00	43,75	62,50	75,00	55,00
Percentil 25°	65,00	68,75	50,00	40,00	56,25	75,00	100,00	80,00
Mediana	85,00	87,50	70,00	55,00	65,63	100,00	100,00	85,00
Percentil 75°	95,00	100,00	80,00	60,00	75,00	100,00	100,00	90,00
Percentil 95°	95,00	100,00	100,00	75,00	81,25	100,00	100,00	95,00
Desv. Típica	22,00	22,13	19,32	15,63	12,38	16,29	9,85	12,13
Mínimo	10,00	12,50	10,00	25,00	37,50	37,50	50,00	45,00
Máximo	95,00	100,00	100,00	85,00	93,75	100,00	100,00	100,00

Tabla 43. Puntuación en las escalas del SF-36, submuestra masculina 75 años y más.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	63,68	71,05	60,00	44,74	55,59	80,26	86,40	77,37
Observaciones	19	19	19	19	19	19	19	19
Percentil 5°	0,00	25,00	10,00	5,00	12,50	25,00	25,00	55,00
Percentil 25°	35,00	43,75	50,00	30,00	50,00	62,50	75,00	65,00
Mediana	80,00	75,00	60,00	50,00	56,25	87,50	100,00	80,00
Percentil 75°	90,00	100,00	70,00	65,00	62,50	100,00	100,00	90,00
Percentil 95°	95,00	100,00	100,00	80,00	87,50	100,00	100,00	100,00
Desv. Típica	31,79	28,81	22,11	21,89	18,62	24,41	25,34	13,78
Mínimo	0,00	25,00	10,00	5,00	12,50	25,00	25,00	55,00
Máximo	95,00	100,00	100,00	80,00	87,50	100,00	100,00	100,00

Tabla 44. Puntuación en las escalas del SF-36, submuestra femenina menores de 25 años

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	97,59	94,91	81,73	74,01	74,54	91,36	91,67	79,01
Observaciones	81	81	81	81	81	81	81	81
Percentil 5°	90,00	75,00	50,00	45,00	43,75	62,50	58,33	55,00
Percentil 25°	95,00	93,75	70,00	65,00	62,50	87,50	83,33	70,00
Mediana	100,00	100,00	80,00	75,00	75,00	100,00	100,00	80,00
Percentil 75°	100,00	100,00	100,00	85,00	87,50	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Desv. Típica	4,27	10,18	21,32	16,89	17,95	15,39	15,48	15,17
Mínimo	80,00	56,25	10,00	25,00	25,00	37,50	25,00	25,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 45. Puntuación en las escalas del SF-36, submuestra femenina entre 25 y 34 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	96,70	95,05	82,78	73,13	73,59	93,26	94,71	79,39
Observaciones	115	115	115	115	115	115	115	115
Percentil 5°	85,00	75,00	50,00	45,00	50,00	62,50	75,00	55,00
Percentil 25°	95,00	100,00	70,00	60,00	62,50	87,50	91,67	75,00
Mediana	100,00	100,00	80,00	75,00	75,00	100,00	100,00	80,00
Percentil 75°	100,00	100,00	100,00	85,00	87,50	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	100,00	93,75	100,00	100,00	100,00
Desv. Típica	5,99	10,55	19,04	16,08	14,46	12,96	10,66	13,20
Mínimo	70,00	43,75	20,00	15,00	31,25	37,50	41,67	35,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 46. Puntuación en las escalas del SF-36, submuestra femenina entre 35 y 44 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	94,04	89,53	78,55	72,23	68,98	90,51	91,97	78,67
Observaciones	83	83	83	83	83	83	83	83
Percentil 5°	75,00	50,00	40,00	30,00	37,50	50,00	50,00	45,00
Percentil 25°	95,00	81,25	60,00	65,00	56,25	87,50	91,67	70,00
Mediana	95,00	100,00	80,00	75,00	68,75	100,00	100,00	80,00
Percentil 75°	100,00	100,00	100,00	85,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	95,00	93,75	100,00	100,00	100,00
Desv. Típica	10,46	16,54	21,02	17,74	18,32	16,75	16,69	17,79
Mínimo	40,00	31,25	20,00	20,00	0,00	25,00	33,33	10,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 47. Puntuación en las escalas del SF-36, submuestra femenina entre 45 y 54 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	91,02	89,38	73,10	64,47	66,43	89,38	92,33	72,26
Observaciones	113	113	113	113	113	113	113	113
Percentil 5°	75,00	50,00	40,00	40,00	37,50	50,00	50,00	40,00
Percentil 25°	90,00	81,25	60,00	55,00	56,25	75,00	91,67	65,00
Mediana	95,00	100,00	70,00	65,00	68,75	100,00	100,00	75,00
Percentil 75°	95,00	100,00	80,00	75,00	75,00	100,00	100,00	85,00
Percentil 95°	100,00	100,00	100,00	85,00	93,75	100,00	100,00	95,00
Desv. Típica	9,33	15,60	19,96	14,54	15,61	16,72	14,99	17,11
Mínimo	40,00	37,50	10,00	20,00	12,50	37,50	25,00	20,00
Máximo	100,00	100,00	100,00	95,00	100,00	100,00	100,00	100,00

Tabla 48. Puntuación en las escalas del SF-36, submuestra femenina entre 55 y 64 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	83,31	83,36	66,49	55,95	62,50	86,32	87,39	76,22
Observaciones	74	74	74	74	74	74	74	74
Percentil 5°	35,00	50,00	20,00	20,00	25,00	37,50	50,00	45,00
Percentil 25°	80,00	75,00	60,00	50,00	50,00	75,00	75,00	65,00
Mediana	90,00	87,50	70,00	60,00	62,50	100,00	100,00	80,00
Percentil 75°	95,00	100,00	80,00	65,00	75,00	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	75,00	87,50	100,00	100,00	100,00
Desv. Típica	18,42	18,66	20,44	15,48	17,22	20,91	19,86	18,32
Mínimo	10,00	25,00	0,00	15,00	18,75	0,00	25,00	20,00
Máximo	100,00	100,00	100,00	85,00	100,00	100,00	100,00	100,00

Tabla 49. Puntuación en las escalas del SF-36, submuestra femenina entre 65 y 74 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	71,22	70,42	59,56	50,33	60,56	83,89	91,48	78,44
Observaciones	45	45	45	45	45	45	45	45
Percentil 5°	40,00	37,50	20,00	25,00	31,25	50,00	50,00	40,00
Percentil 25°	60,00	50,00	40,00	35,00	50,00	75,00	91,67	70,00
Mediana	75,00	68,75	60,00	55,00	62,50	87,50	100,00	85,00
Percentil 75°	85,00	93,75	80,00	65,00	68,75	100,00	100,00	90,00
Percentil 95°	95,00	100,00	100,00	80,00	81,25	100,00	100,00	100,00
Desv. Típica	18,41	22,59	23,25	18,01	17,16	19,15	16,81	17,80
Mínimo	30,00	25,00	20,00	15,00	18,75	25,00	25,00	30,00
Máximo	100,00	100,00	100,00	80,00	100,00	100,00	100,00	100,00

Tabla 50. Puntuación en las escalas del SF-36, submuestra femenina 75 y más años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	53,45	61,64	53,45	41,55	59,48	71,12	89,66	73,45
Observaciones	29	29	29	29	29	29	29	29
Percentil 5°	15,00	12,50	10,00	20,00	37,50	25,00	41,67	45,00
Percentil 25°	30,00	50,00	40,00	30,00	50,00	62,50	100,00	60,00
Mediana	50,00	56,25	60,00	40,00	62,50	75,00	100,00	80,00
Percentil 75°	75,00	93,75	70,00	55,00	68,75	87,50	100,00	85,00
Percentil 95°	85,00	100,00	80,00	70,00	75,00	100,00	100,00	95,00
Desv. Típica	23,38	28,92	22,72	16,86	14,43	25,02	21,09	17,63
Mínimo	15,00	0,00	10,00	10,00	18,75	12,50	25,00	30,00
Máximo	85,00	100,00	100,00	80,00	81,25	100,00	100,00	95,00

Tabla 51. Puntuación en las escalas del SF-36, submuestra sin estudios.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	69,82	71,99	60,63	50,13	61,44	79,46	88,54	77,05
Observaciones	112	112	112	112	112	112	112	112
Percentil 5°	15,00	25,00	20,00	20,00	25,00	25,00	41,67	35,00
Percentil 25°	55,00	50,00	50,00	35,00	50,00	62,50	75,00	67,50
Mediana	75,00	75,00	60,00	50,00	62,50	87,50	100,00	80,00
Percentil 75°	90,00	100,00	80,00	65,00	75,00	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	80,00	87,50	100,00	100,00	100,00
Desv. Típica	25,72	26,41	22,83	20,64	18,68	22,75	20,84	18,69
Mínimo	0,00	0,00	0,00	5,00	6,25	12,50	0,00	15,00
Máximo	100,00	100,00	100,00	100,00	93,75	100,00	100,00	100,00

Tabla 52. Puntuación en las escalas del SF-36, submuestra con estudios primarios.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	87,29	84,99	72,85	62,58	68,02	88,48	91,60	79,39
Observaciones	256	256	256	256	256	256	256	256
Percentil 5°	45,00	37,50	40,00	25,00	37,50	50,00	50,00	45,00
Percentil 25°	85,00	75,00	60,00	50,00	56,25	75,00	95,83	75,00
Mediana	95,00	93,75	70,00	65,00	68,75	100,00	100,00	80,00
Percentil 75°	100,00	100,00	100,00	75,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	95,00	93,75	100,00	100,00	100,00
Desv. Típica	17,13	21,13	22,48	19,72	16,52	18,43	17,23	15,67
Mínimo	10,00	0,00	0,00	5,00	12,50	0,00	25,00	20,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 53. Puntuación en las escalas del SF-36, submuestra con estudios secundarios.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	95,56	93,45	80,63	71,68	72,62	92,95	94,29	80,38
Observaciones	381	381	381	381	381	381	381	381
Percentil 5°	80,00	68,75	40,00	45,00	43,75	62,50	66,67	55,00
Percentil 25°	95,00	93,75	70,00	60,00	62,50	87,50	100,00	75,00
Mediana	100,00	100,00	80,00	70,00	75,00	100,00	100,00	85,00
Percentil 75°	100,00	100,00	100,00	80,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Desv. Típica	8,22	14,59	19,86	16,44	16,06	15,20	14,24	14,60
Mínimo	30,00	0,00	10,00	15,00	25,00	0,00	0,00	15,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 54. Puntuación en las escalas del SF-36, submuestra con estudios superiores.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	94,49	92,16	79,15	69,67	72,04	93,66	95,17	79,21
Observaciones	331	331	331	331	331	331	331	331
Percentil 5º	75,00	56,25	40,00	35,00	43,75	62,50	75,00	55,00
Percentil 25º	95,00	87,50	70,00	60,00	62,50	87,50	91,67	70,00
Mediana	100,00	100,00	80,00	70,00	75,00	100,00	100,00	80,00
Percentil 75º	100,00	100,00	100,00	80,00	81,25	100,00	100,00	90,00
Percentil 95º	100,00	100,00	100,00	95,00	100,00	100,00	100,00	100,00
Desv. Típica	9,64	14,70	20,73	17,57	15,87	12,78	10,04	14,51
Mínimo	35,00	12,50	10,00	15,00	0,00	37,50	50,00	10,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 55. Puntuación en las escalas del SF-36, submuestra con ingresos mensuales inferiores a 900 euros.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	79,49	80,77	67,95	57,76	62,58	83,33	88,03	72,69
Observaciones	78	78	78	78	78	78	78	78
Percentil 5º	30,00	25,00	20,00	20,00	31,25	25,00	41,67	30,00
Percentil 25º	65,00	56,25	50,00	40,00	50,00	75,00	83,33	65,00
Mediana	90,00	100,00	70,00	62,50	68,75	100,00	100,00	80,00
Percentil 75º	100,00	100,00	80,00	75,00	75,00	100,00	100,00	85,00
Percentil 95º	100,00	100,00	100,00	90,00	87,50	100,00	100,00	95,00
Desv. Típica	24,70	26,15	23,43	21,42	16,99	24,48	21,35	19,10
Mínimo	0,00	0,00	0,00	15,00	18,75	0,00	0,00	15,00
Máximo	100,00	100,00	100,00	100,00	93,75	100,00	100,00	100,00

Tabla 56. Puntuación en las escalas del SF-36, submuestra con ingresos mensuales entre 901 y 1500 euros.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	84,86	84,08	72,18	64,30	68,30	87,99	92,13	78,94
Observaciones	179	179	179	179	179	179	179	179
Percentil 5º	35,00	31,25	30,00	25,00	31,25	37,50	50,00	45,00
Percentil 25º	80,00	75,00	60,00	55,00	56,25	75,00	100,00	70,00
Mediana	95,00	100,00	70,00	65,00	68,75	100,00	100,00	85,00
Percentil 75º	100,00	100,00	100,00	80,00	81,25	100,00	100,00	90,00
Percentil 95º	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Desv. Típica	21,33	22,78	23,59	21,93	19,39	20,11	18,38	16,98
Mínimo	0,00	0,00	0,00	5,00	0,00	0,00	0,00	10,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 57. Puntuación en las escalas del SF-36, submuestra con ingresos mensuales entre 1501 y 2000 euros.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	92,18	90,42	79,48	68,47	70,35	92,13	94,18	80,91
Observaciones	308	308	308	308	308	308	308	308
Percentil 5º	65,00	50,00	40,00	40,00	43,75	62,50	58,33	55,00
Percentil 25º	90,00	87,50	70,00	55,00	62,50	87,50	100,00	75,00
Mediana	95,00	100,00	80,00	70,00	68,75	100,00	100,00	85,00
Percentil 75º	100,00	100,00	100,00	80,00	81,25	100,00	100,00	90,00
Percentil 95º	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Desv. Típica	13,77	16,78	20,81	17,68	16,03	13,61	13,86	14,31
Mínimo	10,00	0,00	0,00	15,00	18,75	37,50	0,00	30,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 58. Puntuación en las escalas del SF-36, submuestra con ingresos mensuales entre 2001 y 3000 euros.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	93,06	90,36	76,83	67,48	71,67	92,59	94,17	80,64
Observaciones	319	319	319	319	319	319	319	319
Percentil 5º	75,00	50,00	40,00	30,00	43,75	50,00	66,67	55,00
Percentil 25º	90,00	87,50	60,00	55,00	62,50	87,50	100,00	75,00
Mediana	95,00	100,00	80,00	70,00	75,00	100,00	100,00	80,00
Percentil 75º	100,00	100,00	100,00	80,00	81,25	100,00	100,00	90,00
Percentil 95º	100,00	100,00	100,00	95,00	100,00	100,00	100,00	100,00
Desv. Típica	11,63	17,61	20,72	18,93	15,93	14,85	12,09	14,04
Mínimo	25,00	0,00	10,00	10,00	18,75	25,00	25,00	20,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 59. Puntuación en las escalas del SF-36, submuestra con ingresos mensuales superiores a 3000 euros.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	93,80	91,36	77,30	68,24	72,29	90,82	93,79	78,34
Observaciones	196	196	196	196	196	196	196	196
Percentil 5º	70,00	50,00	30,00	35,00	43,75	50,00	58,33	50,00
Percentil 25º	90,00	87,50	60,00	57,50	62,50	87,50	91,67	70,00
Mediana	95,00	100,00	80,00	70,00	75,00	100,00	100,00	80,00
Percentil 75º	100,00	100,00	100,00	80,00	81,25	100,00	100,00	90,00
Percentil 95º	100,00	100,00	100,00	95,00	93,75	100,00	100,00	95,00
Desv. Típica	10,31	15,84	22,20	17,34	15,56	16,56	13,49	14,84
Mínimo	40,00	12,50	10,00	5,00	12,50	25,00	25,00	20,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 60. Puntuación en las escalas del SF-36, submuestra sin seguro médico privado.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	90,11	88,72	76,40	67,08	69,82	90,93	93,44	79,91
Observaciones	870	870	870	870	870	870	870	870
Percentil 5°	50,00	50,00	40,00	30,00	37,50	50,00	58,33	50,00
Percentil 25°	90,00	81,25	60,00	55,00	62,50	87,50	100,00	75,00
Mediana	95,00	100,00	80,00	70,00	68,75	100,00	100,00	80,00
Percentil 75°	100,00	100,00	100,00	80,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Desv. Típica	16,50	18,91	21,41	19,03	16,57	16,62	14,69	14,84
Mínimo	0,00	0,00	0,00	5,00	6,25	0,00	0,00	15,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 61. Puntuación en las escalas del SF-36, submuestra con seguro médico privado.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	92,62	89,26	75,67	64,98	71,73	89,76	92,86	77,50
Observaciones	210	210	210	210	210	210	210	210
Percentil 5°	70,00	50,00	30,00	25,00	43,75	50,00	58,33	45,00
Percentil 25°	90,00	87,50	60,00	50,00	62,50	87,50	91,67	70,00
Mediana	95,00	100,00	80,00	70,00	75,00	100,00	100,00	80,00
Percentil 75°	100,00	100,00	100,00	80,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	95,00	93,75	100,00	100,00	100,00
Desv. Típica	12,48	19,57	23,98	19,94	17,41	17,74	15,68	17,02
Mínimo	0,00	0,00	0,00	15,00	0,00	25,00	0,00	10,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 62. Puntuación en las escalas del SF-36, submuestra de solteros.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	97,00	94,75	83,32	74,10	74,58	94,16	94,41	81,03
Observaciones	368	368	368	368	368	368	368	368
Percentil 5°	85,00	68,75	50,00	40,00	43,75	75,00	75,00	55,00
Percentil 25°	95,00	100,00	70,00	65,00	62,50	100,00	91,67	75,00
Mediana	100,00	100,00	80,00	75,00	75,00	100,00	100,00	85,00
Percentil 75°	100,00	100,00	100,00	85,00	87,50	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Desv. Típica	8,33	13,56	20,40	17,42	16,49	12,85	12,02	14,28
Mínimo	10,00	0,00	0,00	15,00	0,00	25,00	25,00	10,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 63. Puntuación en las escalas del SF-36, submuestra de casados

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	88,84	86,92	73,67	63,98	68,91	89,77	93,54	79,53
Observaciones	645	645	645	645	645	645	645	645
Percentil 5°	55,00	50,00	40,00	30,00	43,75	50,00	58,33	50,00
Percentil 25°	85,00	75,00	60,00	55,00	62,50	87,50	100,00	75,00
Mediana	95,00	100,00	80,00	65,00	68,75	100,00	100,00	80,00
Percentil 75°	100,00	100,00	100,00	75,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	90,00	93,75	100,00	100,00	100,00
Desv. Típica	15,92	19,75	21,65	18,47	15,82	17,87	15,34	15,12
Mínimo	0,00	0,00	0,00	5,00	6,25	0,00	0,00	15,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 64. Puntuación en las escalas del SF-36, submuestra de viudos, separados y divorciados.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	72,46	74,63	62,39	51,72	58,40	80,78	85,32	69,93
Observaciones	67	67	67	67	67	67	67	67
Percentil 5°	25,00	25,00	30,00	25,00	18,75	50,00	50,00	30,00
Percentil 25°	55,00	50,00	50,00	35,00	50,00	62,50	75,00	60,00
Mediana	80,00	81,25	60,00	55,00	62,50	87,50	100,00	75,00
Percentil 75°	95,00	100,00	80,00	65,00	68,75	100,00	100,00	85,00
Percentil 95°	100,00	100,00	100,00	85,00	87,50	100,00	100,00	95,00
Desv. Típica	25,34	25,65	20,82	20,39	19,17	20,68	21,23	18,98
Mínimo	15,00	0,00	20,00	5,00	12,50	25,00	25,00	20,00
Máximo	100,00	100,00	100,00	90,00	100,00	100,00	100,00	100,00

Tabla 65. Puntuación en las escalas del SF-36, submuestra de no fumadores.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	89,67	88,21	75,79	66,34	69,78	90,31	93,41	79,74
Observaciones	779	779	779	779	779	779	779	779
Percentil 5°	55,00	50,00	40,00	30,00	43,75	50,00	58,33	50,00
Percentil 25°	90,00	81,25	60,00	55,00	62,50	87,50	100,00	75,00
Mediana	95,00	100,00	80,00	70,00	68,75	100,00	100,00	85,00
Percentil 75°	100,00	100,00	100,00	80,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
Desv. Típica	16,24	19,30	21,98	19,58	16,53	17,04	14,51	15,04
Mínimo	0,00	0,00	0,00	5,00	6,25	0,00	0,00	15,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 66. Puntuación en las escalas del SF-36, submuestra de fumadores.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental
Media	93,02	90,41	77,48	67,52	71,24	91,74	93,11	78,67
Observaciones	301	301	301	301	301	301	301	301
Percentil 5°	65,00	50,00	40,00	35,00	37,50	50,00	58,33	50,00
Percentil 25°	95,00	87,50	60,00	55,00	62,50	87,50	100,00	70,00
Mediana	100,00	100,00	80,00	70,00	75,00	100,00	100,00	80,00
Percentil 75°	100,00	100,00	100,00	80,00	81,25	100,00	100,00	90,00
Percentil 95°	100,00	100,00	100,00	95,00	100,00	100,00	100,00	100,00
Desv. Típica	14,45	18,26	21,76	18,26	17,28	16,29	15,83	15,97
Mínimo	0,00	0,00	0,00	5,00	0,00	25,00	0,00	10,00
Máximo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Tabla 67. Puntuación normalizada y componentes del SF-36.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	53,08	52,47	52,09	48,01	55,93	52,79	52,77	52,51	51,40	53,38
Observaciones	1080	1080	1080	1080	1080	1080	1080	1080	1080	1080
Percentil 5°	38,09	37,26	36,76	30,53	41,16	35,03	36,44	35,93	34,27	35,89
Percentil 25°	52,82	49,51	45,21	42,45	52,09	51,40	55,88	48,60	48,13	50,73
Mediana	54,93	56,85	53,67	49,60	58,33	56,85	55,88	52,82	53,48	55,25
Percentil 75°	57,03	56,85	62,12	54,36	61,46	56,85	55,88	58,46	56,86	58,30
Percentil 95°	57,03	56,85	62,12	61,51	70,82	56,85	55,88	64,09	60,21	62,28
Desv. Típica	6,66	7,46	9,27	9,16	8,37	7,35	6,94	8,62	8,12	7,94
Mínimo	14,94	17,67	19,86	18,61	20,87	13,22	9,23	13,40	9,59	9,45
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	72,08	68,59

Tabla 68. Puntuación normalizada y componentes del SF-36, submuestra masculina.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	53,77	52,98	52,95	48,63	56,95	53,62	53,48	53,94	51,87	54,53
Observaciones	540	540	540	540	540	540	540	540	540	540
Percentil 5°	42,30	37,26	36,76	30,53	42,72	40,49	40,33	40,15	35,06	42,95
Percentil 25°	52,82	51,96	49,44	42,45	52,09	51,40	55,88	50,01	49,46	51,71
Mediana	57,03	56,85	53,67	49,60	58,33	56,85	55,88	55,64	54,20	56,00
Percentil 75°	57,03	56,85	62,12	54,36	61,46	56,85	55,88	58,46	57,03	58,74
Percentil 95°	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	60,10	62,46
Desv. Típica	6,26	7,58	9,02	9,30	8,02	6,79	6,36	7,57	7,92	6,89
Mínimo	14,94	17,67	19,86	18,61	23,99	13,22	9,23	16,22	9,59	9,45
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	64,73	67,39

Tabla 69. Puntuación normalizada y componentes del SF-36, submuestra femenina.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	52,38	51,97	51,22	47,39	54,92	51,97	52,05	51,08	50,93	52,23
Observaciones	540	540	540	540	540	540	540	540	540	540
Percentil 5°	35,99	37,26	32,54	30,53	39,60	35,03	32,56	33,11	34,26	32,02
Percentil 25°	50,72	47,06	45,21	42,45	48,97	45,94	51,99	47,19	47,20	49,17
Mediana	54,93	56,85	53,67	49,60	55,21	56,85	55,88	52,82	52,72	54,63
Percentil 75°	57,03	56,85	62,12	54,36	61,46	56,85	55,88	58,46	56,64	57,98
Percentil 95°	57,03	56,85	62,12	61,51	67,70	56,85	55,88	64,09	60,26	62,09
Desv. Típica	6,97	7,31	9,43	8,99	8,58	7,79	7,42	9,35	8,29	8,73
Mínimo	19,15	17,67	19,86	21,00	20,87	13,22	20,89	13,40	12,93	15,77
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	72,08	68,59

Tabla 70. Puntuación normalizada y componentes del SF-36, submuestra menores de 25 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	55,97	54,55	54,63	52,43	58,35	53,44	52,49	53,48	55,17	53,24
Observaciones	171	171	171	171	171	171	171	171	171	171
Percentil 5°	52,82	44,61	36,76	37,68	42,72	40,49	36,44	38,74	44,70	35,96
Percentil 25°	57,03	56,85	49,44	47,21	52,09	51,40	51,99	50,01	52,24	51,31
Mediana	57,03	56,85	53,67	51,98	58,33	56,85	55,88	55,64	56,38	55,35
Percentil 75°	57,03	56,85	62,12	59,13	64,58	56,85	55,88	58,46	59,01	58,30
Percentil 95°	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	61,50	62,28
Desv. Típica	3,58	6,23	9,05	8,14	8,76	6,80	7,45	8,70	5,76	8,74
Mínimo	14,94	17,67	19,86	28,15	33,36	24,13	9,23	16,22	30,65	9,45
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	64,59	62,95

Tabla 71. Puntuación normalizada y componentes del SF-36, submuestra entre 25 y 34 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	55,89	54,58	55,16	50,69	58,13	54,77	53,56	53,37	54,69	53,94
Observaciones	241	241	241	241	241	241	241	241	241	241
Percentil 5°	52,82	47,06	40,99	35,30	45,85	45,94	44,22	41,56	45,08	43,68
Percentil 25°	54,93	54,40	49,44	47,21	52,09	56,85	51,99	50,01	52,91	51,23
Mediana	57,03	56,85	53,67	51,98	58,33	56,85	55,88	55,64	55,88	54,94
Percentil 75°	57,03	56,85	62,12	56,75	61,46	56,85	55,88	58,46	58,01	57,67
Percentil 95°	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	60,13	62,28
Desv. Típica	2,46	5,21	7,80	8,34	6,83	4,55	5,29	6,81	5,15	6,20
Mínimo	40,20	17,67	19,86	23,38	36,48	29,58	9,23	27,48	24,92	23,43
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	63,55	67,39

Tabla 72. Puntuación normalizada y componentes del SF-36, submuestra entre 35 y 44 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	54,76	53,40	53,85	50,66	56,48	53,24	53,18	52,79	53,53	53,23
Observaciones	163	163	163	163	163	163	163	163	163	163
Percentil 5°	46,51	37,26	36,76	35,30	42,72	35,03	36,44	35,93	39,70	37,01
Percentil 25°	54,93	51,96	49,44	47,21	52,09	51,40	55,88	47,19	51,82	50,35
Mediana	57,03	56,85	53,67	51,98	58,33	56,85	55,88	55,64	55,04	55,26
Percentil 75°	57,03	56,85	62,12	56,75	61,46	56,85	55,88	58,46	57,44	58,23
Percentil 95°	57,03	56,85	62,12	61,51	70,82	56,85	55,88	64,09	60,22	62,31
Desv. Típica	4,20	6,80	8,76	8,66	9,09	7,20	6,45	9,35	6,39	8,35
Mínimo	31,78	20,12	24,08	18,61	20,87	24,13	24,78	13,40	22,73	19,58
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	64,73	64,29

Tabla 73. Puntuación normalizada y componentes del SF-36, submuestra entre 45 y 54 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	53,83	52,83	51,84	47,70	55,65	52,83	52,86	51,12	51,97	52,50
Observaciones	220	220	220	220	220	220	220	220	220	220
Percentil 5°	46,51	37,26	36,76	35,30	45,85	35,03	36,44	33,11	40,38	34,36
Percentil 25°	52,82	49,51	45,21	42,45	52,09	51,40	53,94	47,19	48,43	49,91
Mediana	54,93	56,85	53,67	49,60	55,21	56,85	55,88	52,82	52,54	54,49
Percentil 75°	57,03	56,85	62,12	51,98	59,90	56,85	55,88	58,46	56,16	57,52
Percentil 95°	57,03	56,85	62,12	59,13	67,70	56,85	55,88	64,09	60,83	62,16
Desv. Típica	4,07	6,59	8,60	7,37	7,41	7,30	6,76	8,88	6,09	7,93
Mínimo	31,78	17,67	24,08	25,76	27,11	13,22	9,23	19,03	24,16	16,45
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	72,08	65,73

Tabla 74. Puntuación normalizada y componentes del SF-36, submuestra entre 55 y 64 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	50,93	51,40	48,54	44,08	53,67	51,72	51,71	52,11	48,24	52,94
Observaciones	150	150	150	150	150	150	150	150	150	150
Percentil 5°	35,99	37,26	28,31	28,15	36,48	29,58	32,56	33,11	32,03	33,24
Percentil 25°	50,72	47,06	45,21	40,06	45,85	45,94	51,99	47,19	44,55	49,31
Mediana	52,82	54,40	49,44	44,83	55,21	56,85	55,88	52,82	49,44	55,42
Percentil 75°	54,93	56,85	53,67	49,60	61,46	56,85	55,88	58,46	53,20	58,62
Percentil 95°	57,03	56,85	62,12	56,75	67,70	56,85	55,88	64,09	57,60	62,87
Desv. Típica	6,84	7,39	9,05	8,13	8,91	8,46	7,98	9,41	7,21	8,66
Mínimo	19,15	27,47	19,86	18,61	30,24	13,22	20,89	19,03	24,63	19,50
Máximo	57,03	56,85	62,12	61,51	70,82	56,85	55,88	64,09	60,84	66,23

Tabla 75. Puntuación normalizada y componentes del SF-36, submuestra entre 65 y 74 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	45,93	46,86	46,72	40,75	52,06	50,83	53,20	53,28	42,54	55,47
Observaciones	87	87	87	87	87	87	87	87	87	87
Percentil 5°	29,67	32,36	28,31	28,15	39,60	35,03	40,33	33,11	27,65	43,64
Percentil 25°	40,20	37,26	40,99	32,91	45,85	45,94	55,88	50,01	36,54	53,24
Mediana	48,61	47,06	49,44	42,45	52,09	56,85	55,88	55,64	43,26	57,67
Percentil 75°	52,82	54,40	53,67	47,21	58,33	56,85	55,88	58,46	50,10	60,42
Percentil 95°	54,93	56,85	62,12	54,36	61,46	56,85	55,88	64,09	54,55	62,93
Desv. Típica	8,53	8,87	9,18	8,03	7,54	7,80	6,58	8,69	8,87	7,52
Mínimo	19,15	22,57	24,08	23,38	30,24	24,13	20,89	24,66	13,49	22,97
Máximo	57,03	56,85	62,12	56,75	70,82	56,85	55,88	64,09	57,32	64,95

Tabla 76. Puntuación normalizada y componentes del SF-36, submuestra 75 y más años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	39,14	43,28	43,54	36,64	49,81	45,83	50,45	50,01	37,59	53,20
Observaciones	48	48	48	48	48	48	48	48	48	48
Percentil 5°	21,26	22,57	24,08	21,00	36,48	24,13	20,89	33,11	14,56	32,45
Percentil 25°	28,62	36,04	36,76	30,53	45,85	40,49	50,05	44,38	29,23	48,77
Mediana	40,20	45,83	45,21	35,30	50,53	45,94	55,88	52,82	40,03	56,59
Percentil 75°	48,61	54,40	49,44	44,83	55,21	56,85	55,88	58,46	47,08	59,29
Percentil 95°	54,93	56,85	62,12	49,60	61,46	56,85	55,88	61,27	53,29	63,17
Desv. Típica	11,44	11,34	9,50	8,98	8,06	10,88	10,57	9,11	11,78	9,35
Mínimo	14,94	17,67	24,08	18,61	27,11	18,67	20,89	24,66	9,59	31,14
Máximo	54,93	56,85	62,12	54,36	64,58	56,85	55,88	64,09	59,11	68,59

Tabla 77. Puntuación normalizada y componentes del SF-36, submuestra masculina menores de 25 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	55,93	54,27	54,84	53,25	58,58	53,76	52,94	54,58	55,01	54,12
Observaciones	90	90	90	90	90	90	90	90	90	90
Percentil 5°	52,82	32,36	36,76	37,68	42,72	40,49	40,33	41,56	42,88	39,11
Percentil 25°	57,03	56,85	49,44	49,60	55,21	51,40	55,88	52,82	52,77	52,68
Mediana	57,03	56,85	57,89	54,36	58,33	56,85	55,88	55,64	56,67	56,15
Percentil 75°	57,03	56,85	62,12	59,13	64,58	56,85	55,88	61,27	59,01	59,49
Percentil 95°	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	61,18	62,28
Desv. Típica	4,65	7,72	9,14	8,17	8,62	6,89	7,67	8,74	6,28	8,46
Mínimo	14,94	17,67	19,86	30,53	33,36	24,13	9,23	16,22	30,65	9,45
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	62,78	62,95

Tabla 78. Puntuación normalizada y componentes del SF-36, submuestra masculina entre 25 y 34 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	56,11	54,27	55,44	50,33	58,58	55,55	53,69	54,19	54,48	54,68
Observaciones	126	126	126	126	126	126	126	126	126	126
Percentil 5°	52,82	47,06	45,21	28,15	48,97	45,94	44,22	41,56	44,34	46,27
Percentil 25°	57,03	54,40	49,44	47,21	52,09	56,85	55,88	50,01	53,02	51,38
Mediana	57,03	56,85	53,67	51,98	58,33	56,85	55,88	55,64	55,58	55,29
Percentil 75°	57,03	56,85	62,12	56,75	61,46	56,85	55,88	58,46	57,82	57,90
Percentil 95°	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	60,65	62,35
Desv. Típica	2,39	6,03	7,59	8,94	6,45	3,04	5,57	6,10	5,59	5,45
Mínimo	40,20	17,67	19,86	23,38	42,72	45,94	9,23	30,30	24,92	23,43
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	63,12	67,39

Tabla 79. Puntuación normalizada y componentes del SF-36, submuestra masculina entre 35 y 44 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	55,01	54,07	54,67	50,67	57,67	53,78	54,28	53,53	53,79	54,27
Observaciones	80	80	80	80	80	80	80	80	80	80
Percentil 5°	47,56	37,26	36,76	35,30	44,29	37,76	44,22	37,34	38,80	40,51
Percentil 25°	54,93	54,40	49,44	47,21	52,09	56,85	55,88	50,01	53,20	51,70
Mediana	57,03	56,85	53,67	51,98	58,33	56,85	55,88	55,64	55,39	56,07
Percentil 75°	57,03	56,85	62,12	56,75	63,02	56,85	55,88	58,46	57,59	58,77
Percentil 95°	57,03	56,85	62,12	62,71	70,82	56,85	55,88	64,09	60,03	62,31
Desv. Típica	3,98	7,10	8,62	8,91	8,92	7,09	4,49	8,60	6,76	7,41
Mínimo	31,78	20,12	24,08	18,61	23,99	24,13	32,56	16,22	22,73	19,67
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	64,73	62,40

Tabla 80. Puntuación normalizada y componentes del SF-36, submuestra masculina entre 45 y 54 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	54,44	52,99	52,99	48,48	57,34	53,48	53,44	53,93	52,16	54,41
Observaciones	107	107	107	107	107	107	107	107	107	107
Percentil 5°	44,41	37,26	32,54	35,30	45,85	35,03	40,33	41,56	39,18	43,10
Percentil 25°	54,93	51,96	49,44	44,83	52,09	51,40	55,88	50,01	49,51	50,89
Mediana	54,93	56,85	53,67	49,60	58,33	56,85	55,88	55,64	53,35	55,73
Percentil 75°	57,03	56,85	62,12	54,36	61,46	56,85	55,88	58,46	56,34	58,56
Percentil 95°	57,03	56,85	62,12	61,51	67,70	56,85	55,88	64,09	59,92	62,77
Desv. Típica	4,14	7,08	8,65	7,77	6,60	7,29	6,49	7,02	6,44	6,12
Mínimo	33,88	17,67	24,08	30,53	39,60	13,22	9,23	35,93	24,16	34,18
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	63,27	65,73

Tabla 81. Puntuación normalizada y componentes del SF-36, submuestra masculina entre 55 y 64 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	51,83	52,44	49,11	45,24	55,21	52,54	53,37	53,49	48,87	54,51
Observaciones	76	76	76	76	76	76	76	76	76	76
Percentil 5°	38,09	37,26	28,31	32,91	39,60	35,03	36,44	38,74	32,03	35,95
Percentil 25°	50,72	50,73	45,21	40,06	48,97	51,40	55,88	50,01	44,00	50,95
Mediana	52,82	56,85	49,44	44,83	55,21	56,85	55,88	55,64	50,89	56,59
Percentil 75°	54,93	56,85	53,67	50,79	61,46	56,85	55,88	58,46	53,62	58,96
Percentil 95°	57,03	56,85	62,12	59,13	70,82	56,85	55,88	64,09	59,07	63,49
Desv. Típica	5,73	7,37	9,46	8,69	8,99	7,73	6,11	8,28	7,52	7,44
Mínimo	29,67	27,47	24,08	18,61	33,36	24,13	28,67	19,03	27,40	26,57
Máximo	57,03	56,85	62,12	61,51	70,82	56,85	55,88	64,09	60,84	66,23

Tabla 82. Puntuación normalizada y componentes del SF-36, submuestra masculina entre 65 y 74 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	47,01	48,57	48,53	41,31	53,06	51,91	54,58	54,70	43,75	56,65
Observaciones	42	42	42	42	42	42	42	42	42	42
Percentil 5°	27,57	34,81	36,76	30,53	42,72	40,49	44,22	38,74	28,82	45,68
Percentil 25°	42,30	44,61	40,99	35,30	48,97	45,94	55,88	52,82	38,84	54,51
Mediana	50,72	51,96	49,44	42,45	53,65	56,85	55,88	55,64	44,56	57,77
Percentil 75°	54,93	56,85	53,67	44,83	58,33	56,85	55,88	58,46	51,11	60,37
Percentil 95°	54,93	56,85	62,12	51,98	61,46	56,85	55,88	61,27	54,42	63,70
Desv. Típica	9,26	8,67	8,17	7,45	6,19	7,11	4,60	6,83	8,89	5,58
Mínimo	19,15	22,57	24,08	28,15	39,60	29,58	32,56	33,11	13,49	35,28
Máximo	54,93	56,85	62,12	56,75	67,70	56,85	55,88	64,09	57,32	64,95

Tabla 83. Puntuación normalizada y componentes del SF-36, submuestra masculina 75 y más años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	41,75	45,51	45,21	37,55	48,64	48,24	49,54	51,34	40,07	52,77
Observaciones	19	19	19	19	19	19	19	19	19	19
Percentil 5°	14,94	27,47	24,08	18,61	27,11	24,13	20,89	38,74	9,59	31,14
Percentil 25°	29,67	34,81	40,99	30,53	45,85	40,49	44,22	44,38	31,78	47,48
Mediana	48,61	47,06	45,21	40,06	48,97	51,40	55,88	52,82	42,28	53,94
Percentil 75°	52,82	56,85	49,44	47,21	52,09	56,85	55,88	58,46	49,98	59,25
Percentil 95°	54,93	56,85	62,12	54,36	64,58	56,85	55,88	64,09	59,11	65,07
Desv. Típica	13,38	11,29	9,34	10,43	9,30	10,65	11,82	7,76	12,46	9,13
Mínimo	14,94	27,47	24,08	18,61	27,11	24,13	20,89	38,74	9,59	31,14
Máximo	54,93	56,85	62,12	54,36	64,58	56,85	55,88	64,09	59,11	65,07

Tabla 84. Puntuación normalizada y componentes del SF-36, submuestra femenina menores de 25 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	56,02	54,86	54,40	51,51	58,10	53,08	51,99	52,27	55,36	52,27
Observaciones	81	81	81	81	81	81	81	81	81	81
Percentil 5°	52,82	47,06	40,99	37,68	42,72	40,49	36,44	38,74	45,72	34,56
Percentil 25°	54,93	54,40	49,44	47,21	52,09	51,40	48,10	47,19	51,27	50,07
Mediana	57,03	56,85	53,67	51,98	58,33	56,85	55,88	52,82	56,13	53,95
Percentil 75°	57,03	56,85	62,12	56,75	64,58	56,85	55,88	58,46	59,01	58,22
Percentil 95°	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	62,14	62,28
Desv. Típica	1,80	3,99	9,01	8,05	8,96	6,72	7,22	8,55	5,15	9,00
Mínimo	48,61	39,71	24,08	28,15	33,36	29,58	20,89	21,85	43,89	15,77
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	64,59	62,28

Tabla 85. Puntuación normalizada y componentes del SF-36, submuestra femenina entre 25 y 34 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	55,64	54,92	54,84	51,09	57,63	53,91	53,41	52,48	54,92	53,14
Observaciones	115	115	115	115	115	115	115	115	115	115
Percentil 5°	50,72	47,06	40,99	37,68	45,85	40,49	44,22	38,74	46,03	37,90
Percentil 25°	54,93	56,85	49,44	44,83	52,09	51,40	51,99	50,01	52,69	50,96
Mediana	57,03	56,85	53,67	51,98	58,33	56,85	55,88	52,82	55,97	54,91
Percentil 75°	57,03	56,85	62,12	56,75	64,58	56,85	55,88	58,46	58,36	57,46
Percentil 95°	57,03	56,85	62,12	63,90	67,70	56,85	55,88	64,09	59,83	61,02
Desv. Típica	2,52	4,14	8,04	7,66	7,22	5,66	4,97	7,43	4,64	6,85
Mínimo	44,41	34,81	28,31	23,38	36,48	29,58	28,67	27,48	37,56	24,88
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	63,55	63,19

Tabla 86. Puntuación normalizada y componentes del SF-36, submuestra femenina entre 35 y 44 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	54,52	52,75	53,06	50,66	55,33	52,71	52,13	52,08	53,28	52,23
Observaciones	83	83	83	83	83	83	83	83	83	83
Percentil 5°	46,51	37,26	36,76	30,53	39,60	35,03	32,56	33,11	39,70	29,90
Percentil 25°	54,93	49,51	45,21	47,21	48,97	51,40	51,99	47,19	50,55	49,20
Mediana	54,93	56,85	53,67	51,98	55,21	56,85	55,88	52,82	55,02	54,48
Percentil 75°	57,03	56,85	62,12	56,75	61,46	56,85	55,88	58,46	57,14	58,01
Percentil 95°	57,03	56,85	62,12	61,51	67,70	56,85	55,88	64,09	60,34	62,35
Desv. Típica	4,40	6,48	8,88	8,46	9,15	7,31	7,78	10,02	6,04	9,09
Mínimo	31,78	29,91	28,31	25,76	20,87	24,13	24,78	13,40	32,41	19,58
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	63,88	64,29

Tabla 87. Puntuación normalizada y componentes del SF-36, submuestra femenina entre 45 y 54 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	53,25	52,69	50,75	46,96	54,05	52,22	52,30	48,46	51,80	50,70
Observaciones	113	113	113	113	113	113	113	113	113	113
Percentil 5°	46,51	37,26	36,76	35,30	39,60	35,03	32,56	30,30	42,12	29,53
Percentil 25°	52,82	49,51	45,21	42,45	48,97	45,94	51,99	44,38	47,95	47,97
Mediana	54,93	56,85	49,44	47,21	55,21	56,85	55,88	50,01	51,82	52,88
Percentil 75°	54,93	56,85	53,67	51,98	58,33	56,85	55,88	55,64	55,61	56,49
Percentil 95°	57,03	56,85	62,12	56,75	67,70	56,85	55,88	61,27	61,34	60,87
Desv. Típica	3,93	6,11	8,43	6,93	7,80	7,30	6,99	9,64	5,75	8,99
Mínimo	31,78	32,36	24,08	25,76	27,11	29,58	20,89	19,03	38,52	16,45
Máximo	57,03	56,85	62,12	61,51	70,82	56,85	55,88	64,09	72,08	63,62

Tabla 88. Puntuación normalizada y componentes del SF-36, submuestra femenina entre 55 y 64 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	50,01	50,33	47,96	42,90	52,09	50,88	50,00	50,69	47,60	51,33
Observaciones	74	74	74	74	74	74	74	74	74	74
Percentil 5°	29,67	37,26	28,31	25,76	33,36	29,58	32,56	33,11	30,84	32,12
Percentil 25°	48,61	47,06	45,21	40,06	45,85	45,94	44,22	44,38	45,27	47,32
Mediana	52,82	51,96	49,44	44,83	52,09	56,85	55,88	52,82	48,88	54,51
Percentil 75°	54,93	56,85	53,67	47,21	58,33	56,85	55,88	58,46	51,82	58,15
Percentil 95°	57,03	56,85	62,12	51,98	64,58	56,85	55,88	64,09	56,52	62,05
Desv. Típica	7,75	7,31	8,64	7,38	8,60	9,12	9,26	10,32	6,87	9,54
Mínimo	19,15	27,47	19,86	23,38	30,24	13,22	20,89	19,03	24,63	19,50
Máximo	57,03	56,85	62,12	56,75	70,82	56,85	55,88	64,09	58,98	64,52

Tabla 89. Puntuación normalizada y componentes del SF-36, submuestra femenina entre 65 y 74 años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	44,92	45,26	45,03	40,22	51,12	49,82	51,91	51,95	41,40	54,38
Observaciones	45	45	45	45	45	45	45	45	45	45
Percentil 5°	31,78	32,36	28,31	28,15	36,48	35,03	32,56	30,30	27,65	31,70
Percentil 25°	40,20	37,26	36,76	32,91	45,85	45,94	51,99	47,19	34,78	49,84
Mediana	46,51	44,61	45,21	42,45	52,09	51,40	55,88	55,64	40,15	56,49
Percentil 75°	50,72	54,40	53,67	47,21	55,21	56,85	55,88	58,46	49,39	60,45
Percentil 95°	54,93	56,85	62,12	54,36	61,46	56,85	55,88	64,09	54,55	62,79
Desv. Típica	7,75	8,85	9,83	8,58	8,57	8,35	7,84	10,02	8,81	8,89
Mínimo	27,57	27,47	28,31	23,38	30,24	24,13	20,89	24,66	25,26	22,97
Máximo	57,03	56,85	62,12	54,36	70,82	56,85	55,88	64,09	55,92	62,99

Tabla 90. Puntuación normalizada y componentes del SF-36, submuestra femenina 75 y más años.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	37,44	41,82	42,44	36,04	50,58	44,25	51,05	49,13	35,97	53,48
Observaciones	29	29	29	29	29	29	29	29	29	29
Percentil 5°	21,26	22,57	24,08	25,76	39,60	24,13	28,67	33,11	14,56	32,45
Percentil 25°	27,57	37,26	36,76	30,53	45,85	40,49	55,88	41,56	27,65	50,07
Mediana	35,99	39,71	45,21	35,30	52,09	45,94	55,88	52,82	38,58	57,39
Percentil 75°	46,51	54,40	49,44	42,45	55,21	51,40	55,88	55,64	43,60	59,34
Percentil 95°	50,72	56,85	53,67	49,60	58,33	56,85	55,88	61,27	50,61	63,17
Desv. Típica	9,84	11,33	9,60	8,03	7,21	10,92	9,84	9,93	11,24	9,63
Mínimo	21,26	17,67	24,08	21,00	30,24	18,67	20,89	24,66	12,93	31,64
Máximo	50,72	56,85	62,12	54,36	61,46	56,85	55,88	61,27	52,23	68,59

Tabla 91. Puntuación normalizada y componentes del SF-36, submuestra sin estudios.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	44,33	45,88	45,48	40,13	51,56	47,89	50,53	51,16	41,95	53,00
Observaciones	112	112	112	112	112	112	112	112	112	112
Percentil 5°	21,26	27,47	28,31	25,76	33,36	24,13	28,67	27,48	24,34	31,30
Percentil 25°	38,09	37,26	40,99	32,91	45,85	40,49	44,22	45,78	34,69	49,15
Mediana	46,51	47,06	45,21	40,06	52,09	51,40	55,88	52,82	43,03	56,41
Percentil 75°	52,82	56,85	53,67	47,21	58,33	56,85	55,88	58,46	50,52	59,77
Percentil 95°	57,03	56,85	62,12	54,36	64,58	56,85	55,88	64,09	57,93	63,64
Desv. Típica	10,83	10,35	9,65	9,84	9,33	9,93	9,72	10,53	11,23	10,13
Mínimo	14,94	17,67	19,86	18,61	23,99	18,67	9,23	16,22	9,59	19,67
Máximo	57,03	56,85	62,12	63,90	67,70	56,85	55,88	64,09	61,69	68,59

Tabla 92. Puntuación normalizada y componentes del SF-36, submuestra con estudios primarios.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	51,68	50,97	50,65	46,06	54,85	51,82	51,96	52,48	49,48	53,18
Observaciones	256	256	256	256	256	256	256	256	256	256
Percentil 5°	33,88	32,36	36,76	28,15	39,60	35,03	32,56	33,11	31,61	33,00
Percentil 25°	50,72	47,06	45,21	40,06	48,97	45,94	53,94	50,01	45,40	50,20
Mediana	54,93	54,40	49,44	47,21	55,21	56,85	55,88	52,82	51,42	55,38
Percentil 75°	57,03	56,85	62,12	51,98	61,46	56,85	55,88	58,46	55,37	58,78
Percentil 95°	57,03	56,85	62,12	61,51	67,70	56,85	55,88	64,09	59,55	62,69
Desv. Típica	7,21	8,28	9,50	9,40	8,25	8,04	8,04	8,82	8,26	8,66
Mínimo	19,15	17,67	19,86	18,61	27,11	13,22	20,89	19,03	22,73	19,50
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	65,52	67,39

Tabla 93. Puntuación normalizada y componentes del SF-36, submuestra con estudios secundarios.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	55,17	54,29	53,93	50,40	57,15	53,77	53,22	53,04	53,93	53,46
Observaciones	381	381	381	381	381	381	381	381	381	381
Percentil 5°	48,61	44,61	36,76	37,68	42,72	40,49	40,33	38,74	42,60	39,34
Percentil 25°	54,93	54,40	49,44	44,83	52,09	51,40	55,88	50,01	51,44	50,83
Mediana	57,03	56,85	53,67	49,60	58,33	56,85	55,88	55,64	54,86	55,35
Percentil 75°	57,03	56,85	62,12	54,36	61,46	56,85	55,88	58,46	57,82	58,22
Percentil 95°	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	60,52	62,28
Desv. Típica	3,46	5,72	8,39	7,84	8,02	6,63	6,64	8,22	5,65	7,85
Mínimo	27,57	17,67	24,08	23,38	33,36	13,22	9,23	16,22	30,97	9,45
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	72,08	65,73

Tabla 94. Puntuación normalizada y componentes del SF-36, submuestra con estudios superiores.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	54,71	53,78	53,31	49,44	56,85	54,08	53,62	52,38	53,18	53,58
Observaciones	331	331	331	331	331	331	331	331	331	331
Percentil 5°	46,51	39,71	36,76	32,91	42,72	40,49	44,22	38,74	42,07	40,86
Percentil 25°	54,93	51,96	49,44	44,83	52,09	51,40	51,99	47,19	50,34	50,92
Mediana	57,03	56,85	53,67	49,60	58,33	56,85	55,88	52,82	54,75	54,91
Percentil 75°	57,03	56,85	62,12	54,36	61,46	56,85	55,88	58,46	57,60	57,79
Percentil 95°	57,03	56,85	62,12	61,51	70,82	56,85	55,88	64,09	60,53	61,84
Desv. Típica	4,06	5,76	8,76	8,37	7,93	5,57	4,68	8,17	6,31	6,54
Mínimo	29,67	22,57	24,08	23,38	20,87	29,58	32,56	13,40	21,91	24,09
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	63,27	63,49

Tabla 95. Puntuación normalizada y componentes del SF-36, submuestra con ingresos mensuales inferiores a 900 euros.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	48,40	49,32	48,57	43,76	52,13	49,58	50,30	48,71	47,36	50,58
Observaciones	78	78	78	78	78	78	78	78	78	78
Percentil 5°	27,57	27,47	28,31	25,76	36,48	24,13	28,67	24,66	25,92	29,35
Percentil 25°	42,30	39,71	40,99	35,30	45,85	45,94	48,10	44,38	38,99	46,76
Mediana	52,82	56,85	49,44	46,02	55,21	56,85	55,88	52,82	50,14	53,93
Percentil 75°	57,03	56,85	53,67	51,98	58,33	56,85	55,88	55,64	56,28	57,06
Percentil 95°	57,03	56,85	62,12	59,13	64,58	56,85	55,88	61,27	60,98	62,17
Desv. Típica	10,40	10,25	9,90	10,21	8,49	10,68	9,96	10,76	10,71	10,75
Mínimo	14,94	17,67	19,86	23,38	30,24	13,22	9,23	16,22	14,56	9,45
Máximo	57,03	56,85	62,12	63,90	67,70	56,85	55,88	64,09	62,54	68,59

Tabla 96. Puntuación normalizada y componentes del SF-36, submuestra con ingresos mensuales entre 901 y 1500 euros.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	50,66	50,61	50,36	46,88	54,99	51,61	52,21	52,23	49,05	53,43
Observaciones	179	179	179	179	179	179	179	179	179	179
Percentil 5°	29,67	29,91	32,54	28,15	36,48	29,58	32,56	33,11	29,81	32,67
Percentil 25°	48,61	47,06	45,21	42,45	48,97	45,94	55,88	47,19	44,55	49,84
Mediana	54,93	56,85	49,44	47,21	55,21	56,85	55,88	55,64	51,38	56,01
Percentil 75°	57,03	56,85	62,12	54,36	61,46	56,85	55,88	58,46	56,13	59,34
Percentil 95°	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	60,01	62,87
Desv. Típica	8,98	8,93	9,97	10,46	9,69	8,77	8,57	9,57	10,11	8,98
Mínimo	14,94	17,67	19,86	18,61	20,87	13,22	9,23	13,40	9,59	16,45
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	72,08	65,07

Tabla 97. Puntuación normalizada y componentes del SF-36, submuestra con ingresos mensuales entre 1501 y 2000 euros.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	53,74	53,10	53,45	48,87	56,01	53,41	53,17	53,34	52,29	53,77
Observaciones	308	308	308	308	308	308	308	308	308	308
Percentil 5°	42,30	37,26	36,76	35,30	42,72	40,49	36,44	38,74	37,28	38,45
Percentil 25°	52,82	51,96	49,44	42,45	52,09	51,40	55,88	50,01	49,37	51,05
Mediana	54,93	56,85	53,67	49,60	55,21	56,85	55,88	55,64	54,19	55,22
Percentil 75°	57,03	56,85	62,12	54,36	61,46	56,85	55,88	58,46	57,14	58,77
Percentil 95°	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	60,45	62,28
Desv. Típica	5,79	6,57	8,79	8,43	8,01	5,94	6,46	8,06	7,30	7,49
Mínimo	19,15	17,67	19,86	23,38	30,24	29,58	9,23	24,66	24,34	21,79
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	65,52	67,39

Tabla 98. Puntuación normalizada y componentes del SF-36, submuestra con ingresos mensuales entre 2001 y 3000 euros.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	54,11	53,08	52,33	48,39	56,67	53,62	53,16	53,19	52,02	53,94
Observaciones	319	319	319	319	319	319	319	319	319	319
Percentil 5°	46,51	37,26	36,76	30,53	42,72	35,03	40,33	38,74	35,90	42,20
Percentil 25°	52,82	51,96	45,21	42,45	52,09	51,40	55,88	50,01	49,14	51,17
Mediana	54,93	56,85	53,67	49,60	58,33	56,85	55,88	52,82	53,77	54,98
Percentil 75°	57,03	56,85	62,12	54,36	61,46	56,85	55,88	58,46	56,94	58,22
Percentil 95°	57,03	56,85	62,12	61,51	70,82	56,85	55,88	64,09	59,74	62,51
Desv. Típica	4,90	6,90	8,76	9,03	7,96	6,48	5,64	7,90	7,11	7,03
Mínimo	25,47	17,67	24,08	21,00	30,24	24,13	20,89	19,03	21,91	15,77
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	63,55	66,23

Tabla 99. Puntuación normalizada y componentes del SF-36, submuestra con ingresos superiores a 3000 euros.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	54,42	53,47	52,52	48,76	56,98	52,84	52,98	51,89	52,77	52,95
Observaciones	196	196	196	196	196	196	196	196	196	196
Percentil 5°	44,41	37,26	32,54	32,91	42,72	35,03	36,44	35,93	41,50	37,47
Percentil 25°	52,82	51,96	45,21	43,64	52,09	51,40	51,99	47,19	49,65	50,21
Mediana	54,93	56,85	53,67	49,60	58,33	56,85	55,88	52,82	54,34	55,72
Percentil 75°	57,03	56,85	62,12	54,36	61,46	56,85	55,88	58,46	57,70	57,88
Percentil 95°	57,03	56,85	62,12	61,51	67,70	56,85	55,88	61,27	60,28	59,97
Desv. Típica	4,34	6,21	9,38	8,26	7,77	7,22	6,29	8,36	6,68	7,56
Mínimo	31,78	22,57	24,08	18,61	27,11	24,13	20,89	19,03	22,73	19,58
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	64,73	63,54

Tabla 100. Puntuación normalizada y componentes del SF-36, submuestra sin seguro médico privado.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	52,87	52,43	52,15	48,21	55,75	52,89	52,82	52,77	51,30	53,56
Observaciones	870	870	870	870	870	870	870	870	870	870
Percentil 5°	35,99	37,26	36,76	30,53	39,60	35,03	36,44	35,93	33,99	35,97
Percentil 25°	52,82	49,51	45,21	42,45	52,09	51,40	55,88	50,01	48,21	50,83
Mediana	54,93	56,85	53,67	49,60	55,21	56,85	55,88	52,82	53,37	55,22
Percentil 75°	57,03	56,85	62,12	54,36	61,46	56,85	55,88	58,46	56,83	58,37
Percentil 95°	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	60,19	62,28
Desv. Típica	6,94	7,41	9,05	9,07	8,28	7,25	6,85	8,36	8,31	7,78
Mínimo	14,94	17,67	19,86	18,61	23,99	13,22	9,23	16,22	9,59	15,77
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	65,52	68,59

Tabla 101. Puntuación normalizada y componentes del SF-36, submuestra con seguro médico privado.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	53,93	52,64	51,83	47,20	56,70	52,38	52,55	51,42	51,85	52,65
Observaciones	210	210	210	210	210	210	210	210	210	210
Percentil 5°	44,41	37,26	32,54	28,15	42,72	35,03	36,44	33,11	37,56	34,18
Percentil 25°	52,82	51,96	45,21	40,06	52,09	51,40	51,99	47,19	47,49	49,02
Mediana	54,93	56,85	53,67	49,60	58,33	56,85	55,88	52,82	53,89	55,33
Percentil 75°	57,03	56,85	62,12	54,36	61,46	56,85	55,88	58,46	56,94	57,99
Percentil 95°	57,03	56,85	62,12	61,51	67,70	56,85	55,88	64,09	60,71	62,35
Desv. Típica	5,25	7,67	10,13	9,51	8,70	7,74	7,31	9,59	7,29	8,59
Mínimo	14,94	17,67	19,86	23,38	20,87	24,13	9,23	13,40	28,37	9,45
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	72,08	64,52

Tabla 102. Puntuación normalizada y componentes del SF-36, submuestra solteros.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	55,77	54,80	55,07	51,55	58,12	54,30	53,27	53,40	54,95	53,70
Observaciones	368	368	368	368	368	368	368	368	368	368
Percentil 5°	50,72	44,61	40,99	35,30	42,72	45,94	44,22	38,74	45,01	42,58
Percentil 25°	54,93	56,85	49,44	47,21	52,09	56,85	51,99	50,01	52,93	51,21
Mediana	57,03	56,85	53,67	51,98	58,33	56,85	55,88	55,64	56,36	54,92
Percentil 75°	57,03	56,85	62,12	56,75	64,58	56,85	55,88	58,46	58,70	58,25
Percentil 95°	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	60,71	62,28
Desv. Típica	3,51	5,31	8,62	8,31	8,23	5,60	5,61	8,04	5,79	7,32
Mínimo	19,15	17,67	19,86	23,38	20,87	24,13	20,89	13,40	21,91	9,45
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	64,59	67,39

Tabla 103. Puntuación normalizada y componentes del SF-36, submuestra casados.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	52,33	51,73	50,99	46,73	55,30	52,38	52,87	52,56	50,12	53,58
Observaciones	645	645	645	645	645	645	645	645	645	645
Percentil 5°	38,09	37,26	36,76	30,53	42,72	35,03	36,44	35,93	33,99	35,95
Percentil 25°	50,72	47,06	45,21	42,45	52,09	51,40	55,88	50,01	46,41	50,89
Mediana	54,93	56,85	53,67	47,21	55,21	56,85	55,88	52,82	51,89	55,45
Percentil 75°	57,03	56,85	62,12	51,98	61,46	56,85	55,88	58,46	55,49	58,47
Percentil 95°	57,03	56,85	62,12	59,13	67,70	56,85	55,88	64,09	59,44	62,40
Desv. Típica	6,70	7,74	9,15	8,80	7,90	7,80	7,16	8,51	8,00	7,82
Mínimo	14,94	17,67	19,86	18,61	23,99	13,22	9,23	16,22	9,59	19,50
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	65,52	66,23

Tabla 104. Puntuación normalizada y componentes del SF-36, submuestra viudos, separados y divorciados.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	45,44	46,91	46,22	40,88	50,04	48,46	49,03	47,15	44,34	49,73
Observaciones	67	67	67	67	67	67	67	67	67	67
Percentil 5°	25,47	27,47	32,54	28,15	30,24	35,03	32,56	24,66	24,58	29,33
Percentil 25°	38,09	37,26	40,99	32,91	45,85	40,49	44,22	41,56	36,87	45,17
Mediana	48,61	49,51	45,21	42,45	52,09	51,40	55,88	50,01	46,92	53,88
Percentil 75°	54,93	56,85	53,67	47,21	55,21	56,85	55,88	55,64	52,23	57,48
Percentil 95°	57,03	56,85	62,12	56,75	64,58	56,85	55,88	61,27	60,34	62,31
Desv. Típica	10,66	10,05	8,80	9,72	9,58	9,02	9,90	10,69	11,37	11,06
Mínimo	21,26	17,67	28,31	18,61	27,11	24,13	20,89	19,03	14,56	16,45
Máximo	57,03	56,85	62,12	59,13	70,82	56,85	55,88	64,09	72,08	68,59

Tabla 105. Puntuación normalizada y componentes del SF-36, submuestra no fumadores.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	52,68	52,24	51,89	47,85	55,73	52,62	52,80	52,68	51,00	53,53
Observaciones	779	779	779	779	779	779	779	779	779	779
Percentil 5°	38,09	37,26	36,76	30,53	42,72	35,03	36,44	35,93	33,54	35,95
Percentil 25°	52,82	49,51	45,21	42,45	52,09	51,40	55,88	50,01	47,53	50,83
Mediana	54,93	56,85	53,67	49,60	55,21	56,85	55,88	55,64	53,22	55,29
Percentil 75°	57,03	56,85	62,12	54,36	61,46	56,85	55,88	58,46	56,60	58,45
Percentil 95°	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	60,19	62,28
Desv. Típica	6,83	7,56	9,29	9,33	8,26	7,43	6,77	8,47	8,41	7,81
Mínimo	14,94	17,67	19,86	18,61	23,99	13,22	9,23	16,22	9,59	15,77
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	72,08	68,59

Tabla 106. Puntuación normalizada y componentes del SF-36, submuestra fumadores.

	Función Física	Rol Físico	Dolor corporal	Salud General	Vitalidad	Función Social	Rol Emocional	Salud Mental	C. Salud Física	C. Salud Mental
Media	54,10	53,09	52,60	48,42	56,46	53,24	52,66	52,07	52,45	53,00
Observaciones	301	301	301	301	301	301	301	301	301	301
Percentil 5°	42,30	37,26	36,76	32,91	39,60	35,03	36,44	35,93	38,34	35,28
Percentil 25°	54,93	51,96	45,21	42,45	52,09	51,40	55,88	47,19	49,34	50,53
Mediana	57,03	56,85	53,67	49,60	58,33	56,85	55,88	52,82	54,30	55,15
Percentil 75°	57,03	56,85	62,12	54,36	61,46	56,85	55,88	58,46	57,74	58,12
Percentil 95°	57,03	56,85	62,12	61,51	70,82	56,85	55,88	64,09	60,28	62,31
Desv. Típica	6,08	7,16	9,20	8,70	8,63	7,11	7,39	8,99	7,21	8,27
Mínimo	14,94	17,67	19,86	18,61	20,87	24,13	9,23	13,40	22,73	9,45
Máximo	57,03	56,85	62,12	63,90	70,82	56,85	55,88	64,09	63,27	65,73

Apéndice 3. Ejemplo de aplicación del algoritmo SF-6D.

El primer paso necesario para el cálculo de utilidades consiste en traducir las respuestas al cuestionario SF-36 a las seis dimensiones del SF-6D. Para ello hay que seleccionar las preguntas 3a, 3b, 3j, 4c, 5b, 7, 8, 9e, 9b, 9f y 10 del SF-36 (ver Tabla 8 en las páginas 30-31) y traducirlas a los niveles del SF-6D que se describen en la Tabla 6 de la página 28. Por ejemplo, si a la pregunta 8 del SF-36 (“¿Hasta qué punto el dolor le ha dificultado su trabajo habitual?”) el sujeto responde “Bastante”, se le asignará el nivel 5 en la dimensión “Dolor” del SF-6D. Si, por ejemplo, a la pregunta 3a del SF-36 (“Su salud actual, ¿le limita para hacer ... esfuerzos intensos, tales como correr, levantar objetos pesados, o participar en deportes agotadores”) el sujeto responde “No, no me limita nada”, le corresponderá el nivel 1 en la dimensión “Funcionamiento físico”. Y así, sucesivamente.

Una vez hemos traducido las respuestas al SF-36 al sistema SF-6D, pasaremos a aplicar el algoritmo, restando a la unidad los coeficientes correspondientes a los niveles de cada dimensión que caractericen el estado de salud del sujeto. Así, por ejemplo, si la situación de un individuo en un determinado momento del tiempo resultase descrita mediante el estado **231254**, calcularíamos el “peso de calidad de vida” o la utilidad del estado de salud de dicho sujeto como:

$$1 - FF2 - LR3 - DO2 - SM5 - VI4$$

Donde FF2, LR3, DO2, SM5 y VI4 aluden los coeficientes del nivel 2 de “Funcionamiento físico”, el nivel 3 de “Limitaciones del rol”, el nivel 2 de “Dolor”, el nivel 5 de “Salud mental” y el nivel 4 de “Vitalidad”, respectivamente (la dimensión “Funcionamiento social” no supone pérdida alguna de utilidad, pues se encuentra en el nivel 1). Sustituyendo los valores de la columna (2) de la Tabla 22, que se corresponden con el modelo “eficiente” bajo estimación semiparamétrica, tendríamos que la utilidad del estado 231254 es igual a:

$$1 - 0.058 - 0.096 - 0 - 0.044 - 0.203 - 0.124 = \mathbf{0.475}.$$

El algoritmo de Brazier et al. (2002) asigna a este estado una puntuación de 0.681.